

Our best friends are Republicans

That the Republican National Committee has had a hand in choosing members of a scientific advisory committee of the US Department of the Interior threatens the effectiveness of the administration.

So many blacklists of government committee members have appeared in Washington in the past few weeks that nobody whose credentials have been disparaged need be despondent. Indeed, so many honourable people have been blackballed, most notoriously in the Environmental Protection Agency hit-list made public a few weeks ago (see *Nature* 10 March, p.95), that being known as a person of whom the administration approves may be a greater cause of embarrassment. Many will recall Groucho Marx's celebrated remark that he would not wish to belong to a club that would have him as a member. But the revelation that the new membership of a not particularly important advisory committee of the US Department of the Interior has been vetted by officials of the Republican Party raises issues of principle whose importance the administration seems naively to have overlooked. The credibility of the advice it is given on technical questions, if it carries on like this, will undermine the capacity of its successors to govern the United States effectively.

The convention in the United States that a new president appoints not merely cabinet officers but senior members of the civil service from his own political party is well established if imperfectly understood outside the United States. Although the practice has often led to questionable pork-barrel appointments, the staid European opinion that the civil service should rise above party politics is not as easily defensible as it seems. The argument that cabinet officers can be effective only when their policies are supported by their senior civil servants also has force. But hitherto, even US presidents have excluded some parts of the government from the rule that they must be surrounded by their political friends. Thus it has been supposed that the party allegiance (if any) of the director of the National Science Foundation would be irrelevant to his or her tenure of office and, fair play, the Reagan Administration lived peaceably with President Carter's Democratic incumbent until Dr John Slaughter resigned at the end of last year.

What the administration has failed to appreciate is that appointments to advisory committees must fall in a different category. Especially when the subject on which advice is sought is technical, the numbers of people whose expertise qualifies them to give sound advice is almost certain to be limited. Moreover, many of the questions on which the administration seeks advice, often at the insistence of Congress, are strictly above and beyond politics. This is manifestly the case with the Interior Department committee whose membership has now been restricted to registered Republicans. Indeed, to expect that Democrat and Republican geophysicists would differ in their advice on research projects supported by the Department of the Interior is tantamount to supposing that Democrats and Republicans differ from each other not merely on questions such as the role of welfare in a modern society but on theories of the origin of the Earth's crust.

Even when questions with more immediate political implications are to be decided, however, there must be grave doubts whether even the administration's interests are best served by advice prepackaged to give no offence. In the environmental field, for example, this administration is plainly impatient with the network of regulatory legislation that has accumulated, and would have a cogent argument if only it would put its case more temperately (see *Nature* 17 March, p.197). But by no stretch of the imagination can it be supposed that all environmental hazards can

be rendered harmless because the administration wishes that were the case. Dioxin, for example, will remain toxic even in very small amounts, and the administration will have to discharge its legal responsibility for seeing that it is not scattered liberally throughout the United States. Politically, the government will be in serious trouble if it takes too lax a view of how tolerable limits of contamination should be decided. But packing the committees that advise on these questions with yes-men is to court just that risk. A few "snail-darter types", as a now-sacked official of the Environmental Protection Agency called environmentalists, may be politically necessary.

A more sensitive government would have learned this from its experience even in the past few months. When the administration was wrestling last year with the question how and where to base the MX missile, a letter leaked to the newspapers suggested that the chairman of a technical advisory committee, Dr C.H. Townes, disagreed with the Pentagon's dense-pack solution. The administration overrode that unwelcome advice, and then for its pains got egg on its face when it went to Congress for the money. This is the inevitable consequence of ignoring honestly-given advice, whether its authors are Republicans or Democrats. And it will not help to launder committees so as to ensure that the advice they give is always welcome. One result will be to ensure that mistakes are made. Another will be that technical people, usually willing in the national interest to serve presidents for whom they did not vote, will flatly refuse to serve in any capacity. □

Freeing the manipulators

The committees set up to regulate genetic manipulation should be stopped.

DISPENSE with the regulatory committees set up in the past ten years to superintend recombinant-DNA research. That is the simple answer to the central question in the British Government's consultative document published this week (see page 467) — what should be the future of the Genetic Manipulation Advisory Group (GMAG) set up to license recombinant-DNA research projects? The same comment applies to the report *Splicing Life*, put out a few weeks ago in the United States by the President's Commission on Ethical Problems in Medicine (which has itself now gone out of business — see page 468). The regulatory committees, which for a time in the late 1970s were springing up more quickly than the use of restriction enzymes was increasing, have done an invaluable job. But now there is almost nothing left for them to do.

When in 1974 Professor Paul Berg of Stanford University blew the first whistle, pointing out that experiments with the simian virus SV40 loaded with foreign pieces of DNA might endanger people's health, there was good reason why professional people, governments and the general public should have taken fright. The voluntary moratorium on experiments that followed remains without precedent in research. So, too, does the willingness of the research community to fall in with the requirement afterwards imposed that experiments should be specified in advance and carried out only when approved.

This regulatory apparatus was at first designed for two purposes — to protect people's health (that of laboratory workers included) and to assure the general public that progress in this



novel field of research was being made deliberately. The popular nightmare, that genetic manipulators might, even by accident, create especially virulent microorganisms, was never more than a scare: if that were possible in the laboratory, evolution would surely long since have found a way. For a time, however, it was proper that people should be concerned at the movement of oncogenes from one genome to another, or about the possibility that infection with artificial organisms might create immunological problems. With the passage of time, it has become clear that these risks, always hypothetical, are even more remote than they might have been. Now the regulatory committees have become rubber stamps for research proposals in whose assessment laboratory investigators are usually more skilled than most committee members. Many of these proposals are in any case covered by other kinds of legislation. Thus the production of genetically engineered organisms on a commercial scale falls within the scope of legislation on occupational health, work with microorganisms whose natural virulence is a serious threat to human health is covered by legislation of the kind that protects the health of laboratory pathologists and the like, while plant and animal pathogens fall within the scope of regulations designed for the protection of agriculture.

So why not simply abolish the committees, and the requirement that investigators should solicit even *post hoc* approval for proposed experiments? The only justification for keeping this apparatus in being that the public is still disquieted about the potential of genetic manipulation. Can this be so, as the daily newspapers retail how products made by genetic manipulation — one day a sweetening agent, the next a lymphokine — trickle onto the market? The existence of committees such as GMAG in Britain and the Recombinant DNA Advisory Committee (RAC) in the United States has contributed invaluable to this mood. The result is that such residual public anxiety as persists is centred not on questions of safety but, in a loose sense, on what might be called ethics. Under what safeguards and by what criteria will physicians in due course embark on gene therapy — still a somewhat distant prospect (see *Nature* 298, 416; 1982) in spite of some injudicious experiments based on starkly different assumptions. But surely the best safeguard of the public interest against the abuse of clinical techniques is that hospital ethical committees should be effective, and publicly recognized to be so.

All this argues for disbanding both the committees and the procedures that sustain them — a logic that governments seem unwilling to accept. The British Government's preferred solution, that its advisory committee should give advice to the quasi-independent body called the Health and Safety Executive, chiefly concerned with occupational health, is sensible enough: if some part of the government machine senses the need for advice on any subject, that must surely be a sign of grace. But the proposal that the notification procedure should continue as at present, although no longer onerous, is in the circumstances unnecessary, even pointless, and should be dropped.

The President's commission (which was an independent body and which has acted like that) also comes out for some continuing committee but for different reasons. Starting with the proposition that the federal government must have a continuing responsibility not merely for the safety and ethical practice of genetic manipulation but for the promotion of the new technology that follows therefrom, it recommends a standing commission to tackle problems as they arise, giving advice to anyone prepared to listen. Such a commission, if it had the clout, could rid the field of the adversarial politics with which it has been plagued in the past decade. But otherwise the device would be of little value. It is hard to see how a single committee, however well provided with resources, could anticipate and then resolve all the questions, which are certain to arise in this rapidly evolving technology. Modestly, the President's commission says that the proposed standing commission called the "third-generation RAC" should not be a continuation of itself. But it overlooks the dangers in singling out one field of science and technology for special treatment in such a way, and also the present federal government's distaste for new committees. □

Licence for cooperation

May's economic summit may provide a licence for technical collaboration. High time.

Prosperous governments are all alike (and a little smug); every unprosperous government is unprosperous in its own way (and at the same time convinced that it is being done down by others). This is the spirit in which heads of government usually assemble for the meetings now called economic summits which have become a regular feature of Western diplomatic life. But the next such meeting, due to take place next month at Williamsburg, Virginia, will have the benefit of at least one potentially significant emollient of this sense that everybody is at odds with everybody else — the report of the working group on "Technology, growth and employment" commissioned by last year's summit at Versailles and now published (see *Nature* 31 March, p.365). This document should be a reminder to political leaders preoccupied with what they hope is the fag-end of the long recession that the future prosperity of the states they represent will depend critically on their willingness in concert to nurture the technical basis of their domestic economies.

The origins of this document are interesting in themselves, stemming from the case made at Versailles a year ago by the President of France that technology must be an important part of economic introspection. The working party, consisting mostly of chief scientific advisers to governments, has accepted the Mitterrand view that science and technology are the only engines of future prosperity in sight, has deftly devised a set of definitions of the role of governments as sponsors of innovation in the mixed economies of the industrialized West and has made a plea (accompanied by some specific suggestions) for more collaboration between the summit governments on a variety of basic research projects. Artfully, no doubt, the working party has calculated that its document, which is constructive while giving no offence but which is unlikely to be a central part of the agenda, will be politely blessed at Williamsburg in terms that provide officials with responsibility for the administration of research and development with a licence to do what they can.

Not before time. While most of the summit governments pay frequent lip-service to the principle that there should be more international collaboration on basic research, and many of them already collaborate on various projects (CERN, for example, or the Large Space Telescope), they are unreasonably shy of putting their money where their mouths are if there is a hint that collaboration might bring commercial benefits to others than themselves. What else can explain most governments' unwillingness to share the huge cost of carrying fast reactors from the prototype stage to reality — a cause that has been urged on them for at least the past quarter of a century? Why not even go further, as is now suggested, and collaborate on the development of robots, or in biotechnology? For even when the field concerned may yield products that may be commercially exploited in the short term, provided that the rules of commercial competition are agreed — and kept — no collaborator would be left with a sense of having been cheated. To the extent that collaboration is a means of making fuller use of what skill there is, the game is not a zero-sum game.

Even if the Williamsburg summit accepts this proposition only on the nod, we shall all be winners. Even the fear that the more effective use of technical resources by the industrialized nations of the West will further widen the gulf between them and the developing nations may be misplaced, for effective innovation provides the best hope that industrialized states will cease protecting dying industries against competition from developing countries. And if Williamsburg goes according to this plan, there may be a few crumbs for the academic research community if only because everybody now appreciates that even the United States cannot support all the projects, in planetary science for example, that make sense. The summit is unlikely to pledge money for this good course, Italy's and France's wishes apart. But the licence will be sufficient. □

US advisory committees

Party officials accused of screening Interior advisers

Washington

TEN members of a scientific advisory panel of the Department of the Interior were not reappointed last year after they failed to receive "clearance" from the Republican National Committee (RNC). Four other scientists who were cleared by the political group were kept on.

The incident was revealed in a memorandum obtained last week by Senator Dale Bumpers (Democrat, Arkansas) under the Freedom of Information Act. The memorandum, headed "Appointment Clearance Request — Advisory Committee", was sent on 29 January 1982 from the Office of the Secretary of the Interior to RNC. It lists 14 names with blanks after each; the handwritten word "yes" or "no" has been filled in in all but one of the blanks. Only those receiving a "yes" were reappointed (or, in one case, appointed anew) to the advisory committee (the Scientific Committee of the Outer Continental Shelf Advisory Board).

Dr Robert Beardsley of the Woods Hole Oceanographic Institution, a former member of the committee who received a "no" rating, said the group advised the Bureau of Land Management on the scientific merit of research supported by the bureau. "It was not a policy forum", he said. "We never even discussed issues like should there be accelerated lease sales of federal land", one of Secretary James Watt's most controversial policies.

Beardsley said that a meeting of the committee scheduled for early 1982 — the first since the Reagan Administration came in — was cancelled "at the last moment". A couple of months later, he received a form letter saying that his name had been taken off the mailing list.

A spokesman for RNC, Bill Greener, confirmed that RNC had carried out clearance checks on the scientists, and on candidates for other advisory committees and government posts. "What we have done is widespread and extensive", he said. "The only exceptions are Defense, State and Justice."

Greener said "All we do is ascertain that the people are registered (to vote) and are Republicans. We are not passing any judgment whatsoever. A 'yes' means that an individual is registered and a Republican."

Greener said that Carol Williams, the RNC official to whom the memorandum was sent, was a "junior staffer" in the office that handles "White House liaison", Williams, who is now at the Department of Education, refused to comment. The Interior Department official who wrote the memorandum, Derrell Thompson, refused to speak to a reporter who reached him at

his Denver, Colorado, office.

The incident is the latest — and most serious — episode in the continuing story of political manipulation of scientific advisory panels by the Reagan Administration. Congressional investigators looking into allegations of misconduct at the Environmental Protection Agency recently found lists that rated the political leanings of the agency's scientific advisers and senior civil servants. Last year, Food and Drug Administration officials complained that they were being pressed to accept politically inspired choices for posts on independent advisory panels. Also last year, the Department of Agriculture beat a hasty retreat after it was revealed that scientists selected to serve on peer-review panels were being subjected to political checks.

The Interior Department case, however, reveals for the first time the extent of the

role played by RNC, and raises the possibility that many other scientific advisory appointments were similarly scrutinized. "If that group is subjected to political tests, then anybody could be", said James Bruce, a Democratic counsel on the Senate Energy and Natural Resources Committee.

The case also raises the possibility that the law was violated. The charter of the Interior Department committee states that its members are to be selected for their "scientific competence, reputation within their particular fields of expertise, and ability to be representative of important elements of the OCS (Outer Continental Shelf) studies program". The law governing all federal advisory committees requires measures to assure that they "will not be inappropriately influenced by the appointing authority or by an special interest".

But the law does not provide for any criminal penalties, so any question of violation may be somewhat academic. A more critical question, Bruce said, is "What is the long-term impact of this on the willingness of scientists to serve on these panels?"

Stephen Budiansky

Genetic manipulation

Watchdog to bark less often

THE British Government is proposing that its Genetic Manipulation Advisory Committee (GMAG), responsible for licensing experiments in genetic manipulation, should have a less active role but that investigators should continue to disclose experiments they carry out. This proposal is one of four for the future of GMAG contained in a consultative document issued on 5 April and circulated for comment (by 31 May).

The document says that the need to review GMAG's function arises because of "growing indications that early fears about the risks involved in genetic manipulation were exaggerated" (which will seem to many an injudicious phrase), because such concerns as remain about safety centre on large-scale production and because of the reorganization of arrangements for regulating work with dangerous pathogens.

The government is obviously also bothered by the anomalous constitution of GMAG, which by an historical accident has been sponsored since 1976 by the Department of Education and Science under the nose of the then infant Health and Safety Executive and of its supervisory commission, which have direct responsibility for occupational health. The consultative document says that the government departments needing advice no longer include education and science but those responsible for agriculture, environment and industry.

The proposed reorganization, said to have been discussed by GMAG itself, is the favourite among four alternatives —

leaving GMAG alone, splitting it into two with responsibility for advising government departments and the health and safety establishments separately, turning it into a joint committee and, finally, transferring it into a dependant of the Health and Safety Executive as recommended. Presumably the first two options are intended derisively. Abolition is not considered.

Under the proposed reorganization, the committee would be renamed the Health and Safety Commission Advisory Committee on Genetic Manipulation. It would advise the commission and executive on general questions, and give advice to government departments when asked.

The legal requirement that experiments should be notified to the executive would continue, as would the present voluntary undertaking by manufacturers that commercial-scale production plans would be notified. Such data would, however, be adjudicated by officials; only exceptional proposals would go to the committee.

Many researchers will ruefully reflect that the upshot of the proposed reorganization will be that a field of research whose hypothetical hazards are now said to have been "exaggerated" is nevertheless to remain perpetually under regulation. Although for most research laboratories, an annual retrospective report is now sufficient, the principle may stick in some throats.

The government also proposes that the four lay members of the present GMAG should be dispensed with. □

US bioethics

Panel dies a natural death

Washington

THE President's Commission on Ethical Problems in Medicine died a natural death last week, having fulfilled its four-year mandate to investigate half a dozen touchy issues ranging from genetic engineering in humans to compensation for injured research subjects. Despite a long list of pressing ethical problems that the commission did not have time to tackle, and despite considerable interest in the commission's findings by both physicians and the general public, there is for now little hope that Congress will grant the authority needed to revive the body.

The influence of the commission, however, will surely continue to be felt. Besides carrying the *imprimatur* of a President's Commission, the group has, in several very murky areas, become the sole source of clear guidelines. Twelve states and the District of Columbia have already adopted the commission's model definition-of-death statute, which adds "irreversible cessation of all functions of the entire brain, including the brain stem" as an alternative to the conventional definition of cessation of circulatory and respiratory function.

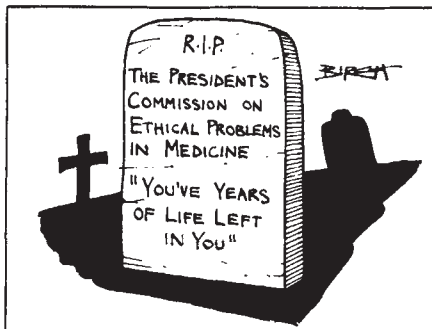
The commission's study of informed consent has been adopted as a text by many medical and nursing schools which, says the commission's director Alex Capron, have felt a need to "get back to basics" — placing more importance on dealing with patients than on mastery of more and more scientific details. And the commission's study makes clear how acute this need is. A scientific poll of physicians' and patients' perceptions, for instance, showed some wide gulfs. While 98 per cent of the doctors said they usually discuss diagnosis and prognosis with their patients, only 78 per cent of the general public agreed. And although 94 per cent of the public said they wanted to be told everything about their condition and treatment no matter how bad the news, only 13 per cent of the doctors said they would give a "straight statistical prognosis" to a patient with a fully confirmed diagnosis of advanced lung cancer, for example. A full 33 per cent said they would say they "couldn't tell how long he might live, but would stress that it could be for a substantial period".

Only last week, the commission released the report that may prove the most far-reaching of all, a study of the ethical problems inherent in decisions to forgo life-sustaining treatment. "For almost any life-threatening condition", the report says, "some intervention can now delay the moment of death. Matters once the province of fate have now become a matter of human choice."

For patients competent to make decisions, the commission comes down squarely on the right of the patient to decide; the

physician's role is to provide the information that will allow that decision to be an informed one.

On the much more difficult question of life-sustaining treatment for incompetent patients, the commission says the decision should rest with a "surrogate" — usually a family member — whose duty it is to act in the patient's best interest. But there must be limitations. For seriously ill newborns, for instance, parents may not make a decision that is "clearly against the infant's



best interests . . . an otherwise healthy Down's Syndrome child whose life is threatened by a surgically correctable complication should receive the surgery because he or she would clearly benefit from it".

The commission recommends that the check in these cases be provided by a hospital "ethics committee" to review

decisions. Above all, the report stresses, hospitals should establish explicit guidelines on how these decisions will be made, and by whom. The commission also says that state legislatures should consider adopting procedures by which people may give advance instructions on their care or grant power of attorney to others to make decisions for them.

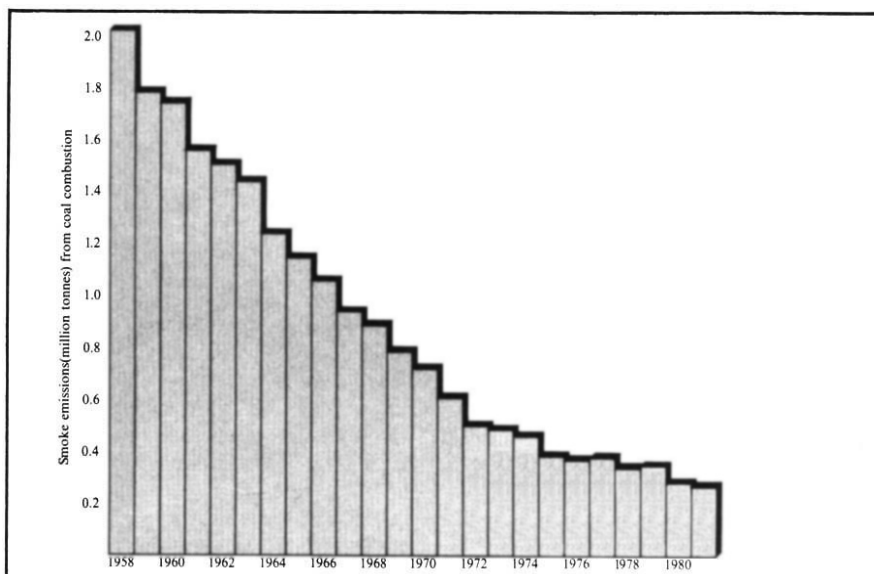
A final report from the commission itemizes some of its unfinished business. A particular thorny problem is whether scientific journals should refuse to publish unethical research: can journals effectively discourage such research by refusing to publish, or is it more instructive to publish questionable research with critical commentary, thereby perhaps preventing its repetition?

Other problems that the commission says need attention "as soon as practicable" include the ethics of diverting drug tests to other countries; of research involving comatose patients; and of representing initial tests of cancer drugs as holding a likelihood of therapeutic value to cancer patients for whom other treatments have failed.

One issue noticeable for its absence from the list — but which Capron said clearly needs the attention of a group such as the commission — is reproductive rights, including a look at the ethics of abortion.

Senator Edward Kennedy (Democrat, Massachusetts) has introduced legislation to renew the commission's authority, but the effort has gained little ground.

Stephen Budiansky



As a result of the 1956 Clean Air Act, the traditional image of a fog-bound London is as up to date as a Trevor Howard movie. The amount of smoke emitted from coal combustion, which, according to the London Weather Centre, was the major cause of the famous "pea-souper" fogs of post-war London, has been reduced by 80 per cent over the past 20 years, and cleaner

fuels such as electricity and gas have taken its place on the domestic market. This trend has been repeated in the electricity industry, where the increased demands for power supply have been met by an increase in the use of oil and, more recently, nuclear power. Source: Digest of Environmental Pollution and Water Statistics, HMSO £6.95.

Leprosy vaccine

Clinical trials begin in Norway

A NEW leprosy vaccine developed under the World Health Organization's Programme for Research and Training in Tropical Diseases has started initial trials with human volunteers in Norway. This marks an important step towards the organization's objective of eradicating leprosy, which now afflicts about 12 million people.

The vaccine was developed at the National Institute for Medical Research in London by a team headed by Dr Richard Rees. Progress towards a vaccine was accelerated in the 1970s with the discovery that the causal agent, *Mycobacterium leprae* — which cannot be grown *in vitro* — could be cultured in armadillos. *M. leprae* grows best at a temperature below that of most mammals, which may explain its preference for the body extremities in humans. Armadillos have a body temperature sufficiently low for *M. leprae* to grow in the liver and spleen.

The new vaccine is based on killed *M. leprae* extracted from armadillos. The crucial step in the development programme — which has taken five years — came with the perfection of a method of inactivating enzymes from the tissues of the host armadillo that previously damaged the bacterial antigens when samples of infected tissue were homogenized. The *M. leprae* are killed by radiation and autoclaved. Tests with laboratory animals show that, surprisingly, the killed bacteria are more immunogenic than live *M. leprae*.

The vaccine is now being manufactured by the Wellcome Foundation Ltd in England, from armadillos kept at the Centre for Applied Microbiology and Research at Porton Down. Clinical trials are expected to start in Britain and the United States before the end of the year. These countries were selected because they have very different BCG tuberculosis immunization programmes, and *M. leprae* is known to interact in some way with BCG. Although the vaccine is intended primarily for use in developing countries, initial trials in developed countries should throw light on safety and on the immune response of the system.

The need for a new weapon against leprosy has been made more urgent as *M. leprae* has evolved resistance to the drug most widely used to treat the disease, dapsone. Other approaches are also being pursued: Dr R Curtiss at Alabama State University is using genetic engineering techniques to produce *M. leprae* antigens. This work is also financed by the World Health Organization.

Dr J Convit, at the Pan-American Centre for Research and Training in Leprosy and Tropical diseases in Caracas, Venezuela, has been using a vaccine similar to that developed in Britain in combination with live BCG tuberculosis vaccine for secondary prevention of lepromatous

leprosy in patients already affected by the disease. Patients with this form of leprosy suffer from a suppression of the immune system, and the BCG/*M. leprae* combination appears to restore immune defences. Early results are said to be promising.

Tim Beardsley

Nucleotide sequences

Who's banker?

Washington

WHEN the National Institutes of Health (NIH) last July awarded a five-year contract worth \$3.14 million to Bolt, Beranek and Newman (BB&N), a private company, to set up a database of nucleic acid sequences, the promise that researchers in genetics would have a central library of sequences seemed to take a major step forward. But now the contract has been challenged by another applicant.

In approving BB&N's proposal, NIH turned down a rival proposal from an existing repository of nucleic acid sequences, the National Biomedical Research Foundation (NBRF), located at Georgetown University here. This is an old and well respected library of both nucleic acid and protein sequences and is contesting NIH's award to BB&N. It will be up to the Comptroller General to decide if NIH followed correct procedures in making their decision.

Meanwhile, NBRF is in trouble. Last March, its founder and principal

investigator, Dr Margaret Dayhoff, died at the age of 57. Dayhoff had started the protein data collection in the early 1960s and in 1965 published the first edition of the *Atlas of Protein Sequences*. A new volume is to be published soon. NIH supported Dayhoff's protein data bank for many years, and still do, but at a low level. And the fate of the databank of nucleic acid sequences is now unclear. Project director Winona Barker says that it is still being updated and has entries of better quality than the BB&N databank. In December it included 1.2 million nucleotides under 862 entries, all annotated as to where the interesting sites are and with other comments. The data are available for a fee.

Meanwhile, the BB&N databank is growing like Topsy. BB&N subcontracted some of the work to the Los Alamos National Laboratory. Los Alamos project director Walter Goad says that it now has 1.37 million bases of which 90 per cent are annotated. More than a hundred subscribers use it through NIH's computer at BB&N's offices in Cambridge, Massachusetts.

The databank incorporates information from the only other collection of comparable size, that at the European Molecular Biology Organization in Heidelberg, West Germany. Both Intelligenetics Inc., a Palo Alto company servicing biotechnology companies, and another data effort at the University of Wisconsin use the Los Alamos data. But if Dayhoff's successors succeed in having the NIH award overturned, this developing pattern will itself be shaken up.

Deborah Shapley

ESA settles for astrophysics

THE European Space Agency (ESA) has selected ISO, an expensive infrared astronomical observatory satellite, for its next scientific mission, despite protests that ESA is neglecting planetary science (see *Nature* 24 February, p.647).

ESA's list of scientific satellites now comprises: Exosat, an X-ray satellite for launch by the US National Aeronautics and Space Administration (NASA) at the end of May this year; Giotto, to fly past Halley's Comet, for launch mid-1985; Hipparcos, for astrometry, for launch towards the end of 1986; and now ISO, for the 'early 1990s'. ESA is also participating in the NASA Space Telescope. The four other proposals before the ESA Science Programme Committee last week, Disco (for solar oscillations), Magellan (UV astronomy), X-80 (X-ray astronomy) and Kepler (a Mars lander), now seem doomed, as another ESA scientific mission will not be selected for another two or three years with no prospect of a launch before 1995. The long gaps between Hipparcos, ISO and the following project are determined partly by the cost of ISO: some 250 million ac-

counting units (AU) (£137.5 million), compared with ESA's science budget of under 100 million AU. Kepler would have cost only 150 million AU.

ISO will be comparable with the currently operating IRAS infrared satellite (a joint UK-Netherlands-US collaboration): it will be cooled with liquid hydrogen and helium to 13 K, and have a lifetime determined by the evaporation rate of the coolants. It will be 100 times more sensitive than IRAS and so able to study extragalactic objects — thus complementing the Space Telescope. Judging from the February report of ESA's Astronomy Working Group, one aim is to give Europe a commanding position in infrared astronomy.

Meanwhile, European planetary science may 'wither away', says Professor Fred Taylor, acting head of the department of atmospheric physics at the University of Oxford and supporter of Kepler. The only remaining planetary mission in view is NASA's Galileo, due to orbit Jupiter in 1989. The only real solution is to increase ESA's mandatory science budget, he says.

Robert Walgate

Polish universities

A legacy of repression

THE period of martial law in Poland — as far as academic life is concerned — was “frequently repressive and always restrictive of universal self-governance”. Such is the summing-up of a 600-page report on human rights in Poland, commissioned by the five-man underground leadership of Solidarity and delivered last month to the heads of national delegations to the Madrid “Helsinki review” conference.

Schools and universities are among those sectors of Polish life singled out for detailed analysis in the report, not only because they are the targets of particularly intensive political “verification” but also because the declaration of martial law came at a time when the universities seemed set on a path of liberalization. This trend, the report claims (citing speeches from the February 1982 Central Committee Plenum), was anathema to the Party, which saw it as a “screen for anti-state indoctrination” and an “extraterritorial enclave of Western institutions of anti-communism”. (The new informal “rector’s conference” was particularly suspect.)

The attack on the universities, says the report, was threefold: from the Ministry of Science, Higher Education and Technology, the Party authorities and the security police.

The ministry was responsible for the removal of university rectors who had been democratically elected under the liberal procedures of 1980–81, the expulsion of student activists from the banned Independent Students’ Association (NZS), the implementation of the political “verification” process (often in the face of concerted passive resistance from academic and administrative staff) and the general disciplining of universities and departments which failed to toe the line. A whole range of sanctions was introduced including the threat of closure of individual departments or whole universities. By the beginning of the academic year 1982–83, twenty of the “democratic” rectors had been removed — including those of Warsaw, Poznan, Wrocław, Gdansk and Torun Universities, the Technical Universities of Krakow, Wrocław, Szczecin, Białystok and Częstochowa, the Higher Naval College of Gdansk and the “Maria Skłodowska-Curie” University of Lublin.

Security police measures included the quelling of demonstrations and also, said the report, gas-grenade attacks on campuses during demonstrations, arrest and interrogation of Solidarity and NZS activists, searches of university premises and the report claims, physical and psychological terrorization of suspects. The case of Dr Jan Hutny of the Wrocław Agricultural College, arrested literally during

the defence of his postdoctoral thesis, is cited as an example of the police policy of attempting to terrorize the academic community.

Many of the examples of repression quoted in the report (which includes some 400 pages of individual reports and testimony) have already been reported. What this report underlines, however, is the illegality of many such acts, notably the notorious Resolution 192 of the Council of State. This made it mandatory for the universities to expel students convicted of offences against martial law (thereby curtailing the competence of the rectors to dis-

cipline their own students) published only on 7 September 1982, it was made retroactive to cover the demonstrations of 31 August.

Many of the most stringent martial law measures — including the temporary return to the draconian 1958 law on the universities which transformed the university rectors virtually into auxiliary police — have been superseded by the new higher education act of May 1982. Other consequences of martial law — including the government ruling that soldiers completing their three-year term of duty can go to university without taking the entrance examination — could well, the compilers of the report fear, become a standard feature of Polish educational policy for years to come. **Vera Rich**

Windscale 1957 accident

Forgotten polonium is traced

It was a difficult week for Mr H.J. Dunster, Director of the UK National Radiological Protection Board (NRPB). The board was forced to admit last week that by an oversight it had omitted to consider a potentially dangerous radioisotope that was released during the Windscale reactor fire in October 1957, in a report published by the board in February this year (see *Nature* 17 March, p.207). The embarrassment is increased by the fact that Mr Dunster contributed to a paper in 1958 that mentioned the isotope, polonium-210. Mr Dunster said last week that the omission “wasn’t too sensible”.

The fire burned for two days in a reactor used for producing materials for nuclear weapons. Polonium was being manufactured by irradiating bismuth, for use with beryllium as a trigger in fission bombs. The fire led to an uncontrolled release into the atmosphere of a gas cloud containing various radionuclides. At the time most of the attention focused on the hazards posed by iodine-131, a gamma emitter that can cause thyroid cancer. However, the board now estimates that about 150 curies of polonium-210 were also released — a fact that appears to have been overlooked by all of the official reports into the incident.

Lord Penney, who chaired the first government inquiry into the affair, said last week that although he had been aware that polonium was in the reactor, it had not been brought to his attention that any polonium was released. A report detailing the level of polonium activity recorded at the Atomic Energy Authority’s base at Harwell appeared in *Nature* in 1958 (182; p.627). Polonium activity in the atmosphere was also detected in Holland and the analysis confirmed at Harwell, where Mr Dunster was at the time health physics advisor for the Windscale plant.

NRPB will be making a full revision of its estimate of population dosage as a result of the fire, but the latest estimates from scientists at the board are that the

polonium would be equivalent to one or two weeks of natural background radiation. This would correspond to an increase of 50 to 100 per cent in the total expected number of “health effects” (cancer deaths plus hereditary effects) — from about 20 to between 30 and 40. This estimate depends on assumptions about the relation between dosage and mortality that are thought to be cautious, so these figures may be thought of as upper limits. Mr J. Moore, Under-secretary of State at the Department of Energy, said in the House of Commons last week that there was no evidence that any member of the public in the United Kingdom had contracted or died from cancer as a result of the fire, and that the inclusion of polonium in NRPB’s estimates did not alter this conclusion.

However, studies published by the Cumbria Area Health Authority do appear to show some increase in the number of registrations of myelomas in southwest Cumbria during the 1970s. Myelomas would not be expected as a result of iodine-131 contamination, but could arise from ingestion of polonium-210, an alpha emitter. Mr Dunster said that the figures — if confirmed — were “odd”, but that it was “inconceivable” that they could be a result of the fire in 1957. John Urquhart of Newcastle University in last week’s *New Scientist* published startling claims that up to 1,000 extra deaths could have been caused by the polonium, but his calculations have been dismissed by NRPB as being based on an out-dated model of polonium transport in the environment, resulting in estimates that are too high by about two orders of magnitude.

Urquhart’s model is based on a NRPB study that has since been revised, but he maintains that the revision is based on evidence relevant to only one of the possible polonium transport routes. NRPB is conducting its own epidemiological study and will be “paying close attention” to the incidence of myelomas. **Tim Beardsley**

US space research

Crying for the Moon

Washington

"THERE'S a feeling that the Moon is going to rise up and bite somebody in the next ten years." This is how one lunar scientist mixed his metaphors in his excitement over the prospects of man's return to the Moon, the subject of the National Aeronautics and Space Administration (NASA)'s fourteenth Lunar and Planetary Science Conference in Houston in mid-March.

Lunar scientists have felt themselves overlooked while other planetary scientists have revelled in the space probe results of recent years. So for two days (while the White House's science office, which has quashed other NASA notions of new space initiatives, was offering a discreet "no comment" to the press) lunar scientists in Houston discussed their possible "giant foothold for mankind", which would come first in the form of modest probes and lead to Moon-based factories fuelling future space transportation systems.

Their argument was that lunar bases and resources may well be useful if the United States, in 20 years or so, decides on a major presence in space. So NASA needs to start building a data base now, by modest efforts such as a polar orbiting lunar satellite to determine whether the polar craters contain ice. If there is water on the Moon, future space missions may be redesigned.

But in their enthusiasm, the lunar scientists almost upstaged another presentation of far more immediate importance to US planetary science. At the meeting, the official Solar System Exploration Committee (SSEC) presented its plan for a low-budget programme of planetary exploration between now and the year 2000 (see below). Among the fourteen initiatives the SSEC has proposed, four have been deemed most urgent. The remaining ten — whose fate is unclear — include the two Moon-related probes.

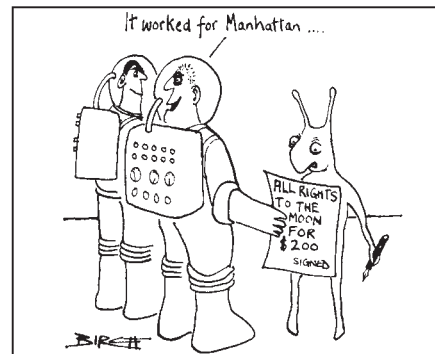
The chairman of SSEC, David Morrison, says he found the ideas for a return to the Moon, both for unmanned probes and manned habitation, "interesting" and "exciting". But a cautionary note was expressed by Philip Chandler of the Office of Technology Assessment of Congress, who was project director for that body's recent report on the poor state of planetary science. Chandler warned that the public was not about to pay the bill for major new US space spectacles and that NASA should think about putting its total programme on a cost-benefit basis. Unless a return to the Moon showed some sign of economic utility, he warned, it would never be approved.

The lunar scientists explained that in 10 or 20 years the United States may wish to mount another large, manned space effort, either for scientific reasons, for interplanetary travel or because of the Soviet

ambitions in space.

"Our original purpose", says W.W. Mendel of NASA's Earth and planetary science group at the Johnson Space Center, "was to make the Office of Space Flight [of NASA] more aware of lunar exploration so they would consider it in their planning". There were also hopes that as part of the follow-on to the shuttle, namely the proposed Space Station and Space Transportation System that would shuttle between the orbiting station and satellites in geosynchronous orbit, a lunar initiative might be added. Lunar enthusiasts also argue that NASA may make decisions, about the fuels for future spacecraft (which could come from the Moon, they argue) or the orbit of the space station, which could rule out a later return to the Moon.

The lunar enthusiasts' near-term programme is modest and many other people seem to think it worthwhile. The first aim is to get the lunar polar orbiter satellite funded, so that the poles, possibly the most useful part of the Moon, can be probed. NASA should also spend about \$1 million per year in Earth-bound studies of possible lunar resources. Finally, there should be a central office to disseminate the information so that it could be taken into account in the main engineering and



scientific programmes.

The more ambitious part of the proposed programme — such as landing men on the Moon again to determine the feasibility of extracting oxygen fuels from lunar materials — is more expensive. But, the lunar enthusiasts argued, such information may be needed for the space transportation system of the future. Even Hans Mark, the NASA associate administrator who gave a keynote address at the meeting, seemed to agree. He argued that man would return to the Moon once "enabling technology" is developed — just as man returned to Antarctica with the development of the airplane as a tool for Antarctic exploration. "I will do whatever I can to speed the day when people will once again put human footprints on the Moon", Mark said in his prepared text.

Deborah Shapley

Ambitious plans for the Solar System

Washington

THE Solar System Exploration Committee (SSEC) presented in Houston its plan for a "core programme" of planetary exploration that it says could be bought for approximately \$300 million per year. This would be half again as much as planetary science's present \$200 million per year — but still a small sum compared with its levels of nearly \$1,000 million in years past. The SSEC plan has met with a favourable response by the White House science office, by some members of Congress and at NASA, according to David Morrison of the Institute for Astronomy of the University of Hawaii, SSEC's chairman. Since SSEC represents many parts of the planetary science community and was formed in 1980 to even out the ups and downs in the programme, its views are likely to be influential.

The core programme would use less state-of-the-art technology and rely on the Mariner Mark II spacecraft adapted for different missions and on modified commercial Earth-orbiting satellites which the SSEC proposes to call the "Observer" class. There are 14 proposals for the next 15 years, of which four get priority.

● Venus Radar Mapper, scheduled for launch in 1988. The committee's early briefings may have caused the Administration to include it in the budget for fiscal

year 1984 (see *Nature* 3 February, p.361).

● Mars geoscience/climatology polar orbiter, to study the climate and map the surface of Mars with infrared, gamma and X-rays.

● Rendezvous with a short-period comet in 1990 using the Mariner spacecraft, which would fly alongside it past the Sun.

● Titan probe to examine the saturnian moon's atmosphere and the possibility of organic processes.

Among the other ten proposals are the polar orbiting lunar satellite (see above), probes to the atmospheres of Saturn and Uranus and a visit to four asteroids.

The \$150 million per year Mariner programme would lend stability and efficiency to planetary science, by frequent launches of common spacecraft, according to Geoffrey A. Briggs, executive director of SSEC.

SSEC is not limiting itself to scientific rationales for exploring the Solar System, according to Briggs. With NASA's blessing, the committee will take another year to study both higher cost, higher profile planetary missions like the Viking programme and the potential for study of resources in near-Earth orbit. And as a concession to the views of the lunar enthusiasts, SSEC recently added a strong advocate of lunar exploration, Larry A. Haskin of Washington University in St Louis.

Deborah Shapley

Spanish "toxic-oil" syndrome

SIR — In their *dual* aetiology hypothesis for the Spanish toxic oil syndrome¹, Root-Bernstein and Westall fail to distinguish between theories and scientific facts. This is precisely one of the major problems we are encountering in Spain when dealing with the several aetiological hypotheses which were put forward in the aftermath of this tragic event.

Infection by *Mycoplasma pneumoniae* was the leading working hypothesis during the first 4–5 weeks, but diagnostic investigations carried out in different laboratories, including the Center for Disease Control in Atlanta in the United States, apparently exclude *M. pneumoniae* as a common infectious agent in the population affected by the toxic syndrome. In fact, only in 2.5 per cent of the investigated cases was *M. pneumoniae* isolated²; furthermore the serological study of 500 paired sera (acute and convalescent) revealed that 2 per cent developed antibodies against *M. pneumoniae*³. There was also no ultrastructural evidence of *Mycoplasma* in post mortem studies of the lung and other organs (M. Rubio, Instituto "Jaime Ferran", CSIC, Madrid, Spain, personal communication). Meanwhile the "toxic-oil" hypothesis was building up on the basis of epidemiological facts and the finding of fatty acid anilides (up to 2,000 p.p.m.) in samples of adulterated cooking oil being used by the affected families.

Fatty acid anilides behave as neurotoxic xenobiotics in rabbits when given orally at a daily dose of 0.01 mg per kg body weight (letter to *Lancet*, in the press) which is one or two orders of magnitude below the expected daily human intake (on body weight basis) from an adulterated oil containing 1,000 p.p.m. fatty acid anilides. In addition, the immunogenicity of fatty acid anilides to rabbits has been established by intradermal challenge in anilide treated animals, solid phase radioimmunoassay of their sera and by the immunofluorescent detection of "anilide specific antigens" in tissue slices of anilide treated animals, incubated with immune serum and fluorescein labelled goat anti-rabbit serum⁴.

Root-Bernstein and Westall suggest that anilides might be acting as "adjuvants" for a hyperimmune response to some mycoplasma. Our experimental findings unambiguously show that anilides alone behave as immunogenic and neurotoxic to experimental animals. A dual hypothesis might not be necessary to explain the Spanish toxic-oil syndrome.

L. VALENCIANO
E. MUNOZ
V. LARRAGA
A. PESTANA

Consejo Superior de Investigaciones Científicas, Madrid, Spain

1. R.S. Root-Bernstein & Westall, F.C. *Nature* **301**, 178 (1983).

2. J.G. Riagau-Perez & Winkler, W.G. *Lancet* **ii**, 724 (1982).
3. Grupo de Investigación Microbiológica, *Boletín Epidemiológico Semanal* (Ministerio de Sanidad y Consumo) **1485**, 153 (1981).
4. A. Marquet *et al.*, presented to the "CSIC workshop on the Toxic-Oil Syndrome", Madrid (Spain), March 14–16, 1983.

Soviet scientists under pressure

SIR — In *Nature* of 24 December 1981 (**294**, 688) you published a letter entitled "Voices from a wilderness". It described the plight of scientists in the Soviet Union who have requested permission to emigrate, including their loss of employment and the revocation of their scientific degrees.

We have recently returned from a visit to the Soviet Union where we met several of the scientists whose names appear on the published correspondence. The purpose of our visit was to deliver offers of visiting professorships to seven of the scientists. Representing the faculty of the University of Southern California, we wanted to express our concern for their rights and academic freedom and to create the opportunity for them to resume research and teaching.

We think that the editor and the readership of *Nature* will want to know of recent events in connection with the December 1981 correspondence. When we were meeting with the Soviet scientists, we learned that the persons whose names appeared on the *Nature* letter were being summoned to the prosecutor's office for intense interrogation. Each was called separately as a "witness" and shown a copy of the *Nature* letter, but with his own name missing from the list of signers.

We were told that one of two goals were intended: one or more of the scientists would agree to be a witness against the others, who would then be prosecuted; or, all the scientists would deny connection with the letter and the authorities could then publicly claim the *Nature* letter was a fabrication of the West. Thus far, we have also been told, those whose names appeared in the *Nature* letter have foiled these intentions. It appeared to us likely they will remain united and continue to refuse to answer the prosecutor's questions.

These scientists have been and are courageous. They face a mounting campaign by the Soviet authorities to intimidate them and their families. A protest by *Nature* and its readership over the letter in *Nature* would, we think, be appropriate.

MICHAEL MELNICK
IRWIN C. LIEB

*University of Southern California,
Los Angeles,
California, USA*

Intersexual selection

SIR — Ornithologists for over a century have been fond of criticizing the Darwin-Fisher concept of female choice as a factor in the evolution of extreme sexual dimorphism. The loudest criticisms seem to result from a lack of proper experiments, coupled with a puzzling resistance to attempting them^{1,2}. Andersson's³ study beautifully illustrates the power of experiments in a natural population to clarify age-old disputes in behavioural ecology. However, the report, like many others in the recent literature on sexual selection, continues a curious usage of a term that does not mean what it says, namely "intersexual selection". According to Darwin⁴, sexual selection depends "on a struggle between the individuals of one sex, generally the males, for the possession of the other sex". Therefore, regardless of whether aggression or choice is prominent, sexual selection must always by its very nature be intrasexual; it can never be intersexual.

Although the term "intersexual selection" was not actually introduced by Huxley⁵, his introduction of "intrasexual selection" for aggressive struggles between males for females strongly implied that the principal alternative, "epigamic selection"⁶, must be "intersexual selection." Certain recent textbooks⁷ have unfortunately preferred Huxley's terms to Darwin's. Some recent authors^{8–10} have simply substituted "intersexual selection" for "epigamic selection". The behavioural interactions between the female and male are, of course, intersexual, but the evolutionary selective process, even in mate choice uncomplicated by male-male aggression, is inherently intrasexual. The males compete among themselves for the female's preference and as a result of female preference one genotype of male may increase relative to another.

For Darwin's distinction between preference and aggression as factors in sexual selection, the terms intra- and intersexual selection are at best misleading. Both should be abandoned. In their place the English language provides adequate material with which to make the distinction between sex-ual selection based on aggression and that based on preference. Darwin did not need jargon for this distinction.

JERRAM L. BROWN

*Department of Biological Sciences,
State University of New York,
Albany, USA*

1. LeCroy, M. *Am. Mus. Novitates* No. 2714 (1981).
2. Payne, R.B. & Payne, K. Z. *Tierpsychol.* **45**, 113–173 (1977).
3. Andersson, M. *Nature* **299**, 818–820 (1982).
4. Darwin, C. *The Origin of Species* (Dutton, New York, 1928).
5. Huxley, J.S. *Am. Nat.* **72**, 416–437 (1938).
6. Poulton, E.B. *The Colours of Animals* (London, 1890).
7. Alcock, J. *Animal Behavior* (Sinauer, Sunderland, Massachusetts, 1979).
8. Wilson, E.O. *Sociobiology* (Harvard University Press, Cambridge).
9. Loiseleur, P.V. *Am. Nat.* **120**, 721–732 (1982).
10. Halliday, T.R. in *Behavioural Ecology* (eds Krebs, J.R. & Davies, N.B.) 180–213 (Blackwell, Oxford, 1978).

What chance a malaria vaccine?

As hopes recede that traditional means of controlling malaria will succeed, the promise and problems, both scientific and political, of malaria vaccines are coming into focus.

VACCINES against malaria are suddenly all the fashion. Research groups, the World Health Organization (WHO) and biotechnology companies seem eager to exploit recombinant DNA and monoclonal antibody techniques to isolate and produce those parasite antigens that might form the basis of vaccines. So it is important to realize that the current interest in vaccines has also been stimulated by the declining effectiveness of earlier methods of control.

The optimism of past decades that a combination of mosquito eradication and treatment of the disease would provide effective control of malaria has now been largely dispelled, several success stories notwithstanding. Partly because of indiscriminate use, these methods are no longer so effective. The problem is resistance — to the insecticides, mainly DDT, used to eradicate mosquitoes (successfully in Italy, Greece and parts of South America) and to chloroquine, once the best of the few drugs active against the disease but fast becoming ineffective as resistance to it has spread from Asia to South America and now to Africa. (Conceivably others will in time be developed on the basis of the newly discovered susceptibility of the malarial parasite to oxygen or peroxide radicals — see *Nature* 3 March, p.19.) With only one new drug, mefloquine, on the horizon the immediate prospects of controlling malaria look bleak and it is spreading back to some areas at one time freed of the disease.

The new enthusiasm for vaccines is thus understandable, particularly with the prospect that biotechnology will facilitate their production. The first success of the biotechnological approach is reported on page 536 of this issue.

The success is a just reward for Dr Ruth Nussenzweig's persistence in her studies of the sporozoite form of the malaria parasite, which is injected into the bloodstream by a mosquito bite. By all accounts, it puts the prospect of a sporozoite vaccine about two years ahead of a vaccine against forms of the parasite arising at later stages in its life-cycle. But, not least because the cloned gene is from a species of *Plasmodium* that causes malaria in monkeys rather than one of the four that infects man, the work is more of a hop than a jump towards an effective vaccine.

In any case, there are doubts that a sporozoite vaccine alone will ever be effective in itself. The problem is that the 1,000 or so sporozoites injected by a mosquito

bite are, like ballistic missiles, difficult targets and apt to cause great damage even if only one gets through. There is only about 30 minutes from their entry into the bloodstream in which to destroy the sporozoites before they go to ground in the liver. Once there, a process of development and replication leads to the production of many thousands of merozoites which are released into the bloodstream to enter red cells, continue multiplication and cause the symptoms of malaria. To a first approximation, the final number of merozoites is the same however many sporozoites enter the liver, so unless a vaccine against sporozoites removed them all, the course of the disease will be slower but the end result the same.

Even so, the sporozoite vaccine may have one practical advantage over the merozoite vaccine which many favour — it may work without the need for an adjuvant, a substance that enhances the antibody response to an antigen with which it is given. In several studies, including one on people, protection against malaria has been induced simply by prior exposure to sporozoites rendered incapable of further development by irradiation. Whether the purified surface antigen of the sporozoite will also work in this way is not yet known.

Merozoites, by contrast, require an adjuvant and none that is available is acceptable for use in people. A different problem confronts those developing the third possible kind of vaccine — that against the gametocyte stage of the parasite which develops from the merozoite and which continues the life cycle when taken back into mosquitoes. For such a vaccine, however effective, would do nothing for those vaccinated since it is aimed only at preventing transmission of the disease.

Over and above these problems hovers a general cloud of doubt about the likely effectiveness of any vaccine against malaria, given the persistence of the disease in the face of circulating antibodies in many patients.

Even so, and despite the practical difficulties in developing countries of vaccinating children in their first year of life — when most malaria infections and fatalities occur — all three vaccines are still worth pursuing. If, however, the sporozoite vaccine is anything to go by, progress will be complicated not only by scientific and technological problems but also by the clash of political and commercial interests.

Malaria is a disease of developing countries influential on WHO which supports the majority of malaria research. Not surprisingly, WHO insists that discoveries stemming from research it sponsors be freely available to its members. That presents difficulties when, as in the case of Dr Nussenzweig's work, the institution in which she is employed takes out patents and then enters into negotiations with a biotechnology company (in this case, Genentech) that demands exclusive rights on any vaccine produced in return for carrying out the further development of the product.

It is not yet clear to what extent a clash between the vested interests of the company and WHO accounted for the delay and eventual breakdown of negotiations in this instance (the company maintains its lost interest for other reasons — see *Nature* 31 March, p.366).

It is certain that delay cannot be in the interests of the 200 million infected with malaria at any one time or the 2 million who die from it each year. Since biotechnology companies are in business to make profits and since there can, as yet, be no assurance that a malaria vaccine will be profitable, it is understandable that a company should drive a hard bargain. It is equally understandable that politicians from developing countries will resent the prospect of Californian entrepreneurs growing rich on their problems. How is that dilemma to be resolved?

The best option for WHO must be to persuade the institution holding patent rights to do no more initially than to contract with a biotechnology company to produce vaccine for initial trials. WHO would have to foot the bill for that and thereafter, if the vaccine looked promising, would probably be best off striking a commercial deal with a company (not necessarily the same) to produce the vaccine, to market it in some countries but also to license it to any other members of WHO. It might also be made a condition of WHO support that it becomes the beneficiary of any patents that may arise. For if there are objections to the prospect of a risk-taking company making a profit from the development of a WHO-supported discovery, surely there are even more objections to the possibility that the institution where any such discovery is made should, without taking any risk, stand to profit at the expense of WHO.

Peter Newmark

Tumours of lymphocytes

What happens when cellular oncogenes collide with immunoglobulin genes

from Miranda Robertson

ABOUT 18 months ago, George Klein suggested in *Nature*¹ that the events leading to malignancy in some B-cell lymphomas might represent some kind of travesty of the genetic rearrangements underlying the normal differentiation of B lymphocytes. In consequence, the past year has seen a remarkably rapid convergence of research on lymphocyte malignancies and that on the immunoglobulin genes that undergo the rearrangements; and last month, representatives of both fields met in Basel* to discuss how far Klein's arresting idea has been supported by the rush of research that has followed its publication.

The hypothesis was originally inspired by the observation that a very high percentage of Burkitt's lymphomas and mouse plasmacytomas have translocations involving chromosome 8 (in man) or chromosome 15 (in mouse) and one of the three chromosomes containing the immunoglobulin genes. Thus Klein suggested that the intrachromosomal breakages that in normal B lymphocytes bring one of the silent variable (V)-region immunoglobulin gene segments into the transcriptionally active site of the constant (C)-region segments are replaced in some tumour cells by chromosomal translocations that instead introduce a cellular oncogene into the same active site. He accordingly predicted that a cellular oncogene should be discovered at the breakpoints on chromosomes 8 and 15, and with the discovery that the *c-myc* gene is indeed located at the breakpoint², it has begun to look as though the theory is in principle right. In detail, however, several complications have become apparent.

Movements of genes

First, it is clear that the activation of *c-myc* is not in any simple way equivalent to the activation of V-region gene segments, which recombine with downstream J segments adjacent to the C-region segments through specific conserved signal sequences for both heavy- and light-chain genes. So far, there are detailed data only for translocations involving the heavy-chain locus, and there is overwhelming evidence that in these cases the breakpoint is at the site of a quite different set of conserved recombination signals — those mediating the heavy-chain switch, down-

stream of J and upstream of each of the heavy-chain C-region genes (S. Cory, Walter and Eliza Hall Institute of Medical Research; K. Marcu, State University of New York at Stonybrook; T. Rabbitts, MRC Cambridge).

This predilection for switch sites cannot, however, be explained as an exploitation of the mechanisms that mediate the normal heavy-chain switch, since there seems to be no homology between the switch sequences and the site of the break on chromosome 15 where the mouse *c-myc* gene lies³. (There is a strongly preferred region of breakage on chromosome 15 (ref. 4) but this can probably be explained by selection for specific effects of the site of the breakpoint in the behaviour of the *c-myc* gene — see later.) And while the switch-site preference may explain why the vast majority of translocations involve the heavy-chain locus rather than either of the two light-chain loci, which do not undergo C-region switching, it leaves the precise significance of the switch sequences in translocation something of a mystery. The fall-back position on this is that normal recombination events at the switch sites predispose to chromosomal breakage at those sites; that this occasionally results in a translocation involving the site of the *c-myc* gene on chromosome 8 or 15; and that the high frequency of such translocations in lymphoma and myeloma cells is due to selection for the effects of such translocations on the *c-myc* gene.

Transcriptional activation

Further complications arise in connection with the transcriptional activation of the translocated *c-myc* genes. In normal B cells, one of several silent V segments is activated by the combined effect of its own upstream promoter and the constitutively active state of the chromatin at the C region with which it recombines. It is now clear that *c-myc* is not activated by this combination.

First, it translocates to the allelically excluded chromosome in which no productive recombination of V (and its promoter) with C segments has taken place; and second, in any case it is translocated in the reverse transcriptional direction from that of the immunoglobulin genes. Thus the V-region promoter can play no part, although the active state of the chromatin at the C region may clearly still be important. Indeed, C. Croce (Wistar Institute, Philadelphia) has conclusively demonstrated, by segregating the translocated and non-translocated *c-myc*

genes from the same lymphoma in different somatic cell hybrids, that transcriptional activity in the translocated gene is increased, at least in this lymphoma.

At the same time, however, it has become clear that the translocated *c-myc* gene may become truncated in transit, with no clear increase in transcription⁵, and this has given rise to the alternative theory that it may be alterations in the structure of the gene, rather than in its regulation, that confer its putative tumorigenic properties. More detailed investigations on the translocated gene and its normal counterparts have now reconciled these possibilities by showing how alterations to the gene could lead to alterations in the level of the gene product.

It now emerges that both in mouse plasmacytomas (M. Cole, St Louis; Cory; Marcu) and in Burkitt's lymphomas (R. Taub, Harvard), an extremely common consequence of the translocation is to break the *c-myc* gene in a large intron separating the first of three exons from the remaining two. At first sight, this raises further problems for the transcriptional activation of the gene, since its normal promoter is presumably lost with the 5' exon. However, Cory, Taub and Cole all report that the translocated portion of the 5' intron contains sequences conserved between mouse, man and chicken; and since Cole and Marcu⁶, who have sequenced some of the mouse gene, find a TATA box and a cap site in the 5' intron, it seems likely that these are what the conserved regions contain, and the translocated truncated *c-myc* gene thus brings a cryptic downstream promoter with it. Moreover, it is only the two 3' exons of the *c-myc* gene that are present in the viral *myc* gene; so it looks very much as though the loss of the 5' exon may be the crucial event in the activation of the transforming activity of the gene. How this loss might contribute to an increase in the dosage of the gene product, even in the absence of any obviously increased transcriptional activity, has been suggested for mouse myeloma cells by Stanton *et al.*⁶.

The answer may lie in the nature of the 5' exon of the intact *c-myc* gene. Although the exon is transcribed in normal cells, it does not code for protein since it contains several stop codons that clearly prohibit its translation. It is possible that the existence of this long untranslated region may greatly reduce the efficiency of translation of the two functional exons, so that its removal

*A meeting on 'Mechanisms of B-cell neoplasia' was held at the Basel Institute of Immunology on 15–17 March 1983. The proceedings of the meeting will be published and a limited number of copies is available free of charge from Ms Ingrid Westen-Schlur at the Basel Institute of Immunology, upon written request.

Miranda Robertson is Associate Editor of *Nature*.

increases translational efficiency and results in the synthesis of more *myc* product with little increase in *myc* transcription.

How far the same scheme could account for the events underlying the transformation of Burkitt's lymphoma cells is not absolutely clear, since in some cases the truncated gene does not seem to be transcribed at all in tumour cells, while in others a truncated transcript is produced from a *c-myc* gene that has been translocated intact. Rabbits, discussing these paradoxes, suggests that the human gene, like the mouse gene, may contain a downstream promoter; and that for some reason the downstream site is preferentially used in the context of the immunoglobulin locus. This could account for the production of truncated transcripts from a complete translocated gene. The absence of the shorter transcript in some cells with truncated *c-myc* genes may simply be a reminder that not all routes to neoplastic transformation lead through *myc*.

Other oncogenes

Thus, as well as there being more than one route to the potentiation of *c-myc*, there is more than *c-myc* potentiation to the production of a tumour cell. Several groups have already reported that although DNA from lymphomas and plasmacytomas with *c-myc* translocations will transform NIH 3T3 cells, the transforming gene in these assays is not *c-myc* (which has never been known to transform this fibroblast line)^{3,7,8}.

M. A. Lane, from Cooper's laboratory (Harvard), was able to report in Basel on the identity of at least one of the non-*myc* transforming genes. She and her colleagues have recently identified the transforming gene from chicken bursal lymphoma cells as a hitherto unknown oncogene with partial homology to transferrin⁹. They have now isolated from Burkitt's lymphoma cells a gene hybridizing with the chicken probe, and shown that it will also transform NIH 3T3 cells. The contributions of other known cellular oncogenes have been the topic of two recent articles in *News and Views*^{10,11}.

Perhaps the next 18 months will see the beginning of the task of tracking the activities of the increasing number of implicated oncogenes through normal ontogeny and in tumours of different types, and the development of a complete picture of the multi-step process of tumorigenesis in lymphoid cells of the B lineage. □

1. Klein, G. *Nature* **294**, 313 (1981).
2. see Newmark, P. *Nature* **301**, 111 (1983).
3. Adams, J. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
4. Cory, S., Gerondakis, S. & Adams, J. *EMBO J.* (in the press).
5. see Newmark, P. *Nature* **301**, 462 (1983).
6. Stanton, L.W., Watt, R. & Marcu, K. *Nature* (in the press).
7. Cooper, G.M. & Neiman, P.E. *Nature* **292**, 857 (1981).
8. Crews, S. *et al. Science* **218**, 1319 (1982).
9. Goubin, G. *et al. Nature* **302**, 114 (1983).
10. see Newmark, P. *Nature* **300**, 582 (1982).
11. see Newmark, P. *Nature* **300**, 682 (1982).

Astronomy

Quasars: bridging the radio to infrared gap

from Alan P. Marscher

WHEN quasi-stellar radio sources (quasars) were first discovered, they could be studied only in the optical and radio wavelengths (and by theoretically oriented minds). Since then, space-borne ultraviolet, X-ray and, to a lesser degree, γ -ray instruments have opened up new regions of the spectrum to quasar investigators. On the ground, large telescopes on high mountaintops have enabled routine observations of quasars to be performed within the atmospheric wavelength 'windows' of the infrared (IR). Only recently, however, have attempts been made to explore the spectral gap between the microwave radio and far-IR: the millimetre and submillimetre wavelengths. These efforts have been rewarded with the acquisition of important information which should help solve some of the mysteries of quasars.

There are difficulties in making millimetre and submillimetre wavelength observations because the Earth's atmosphere is quite opaque to most electromagnetic waves in this portion of the spectrum, and putting one's instruments high up on the tops of extinct volcanoes only partially alleviates the difficulty. Compounding the problem is the impossibility, at present, of building a detector that can yield the high sensitivity required to detect faint astronomical objects. Astronomers are thus forced to observe only relatively bright objects from high-altitude jets or mountaintop observatories. Nonetheless, some quasars are bright enough at millimetre to far-IR wavelengths that they can be studied using currently available instruments.

It is important to measure the flux of millimetre to IR wavelength emission from quasars for several reasons. At centimetric radio wavelength, most 'radio-loud' quasars have spectra which are quite flat. At least for several well studied flat-spectrum quasars, the effect is mostly due to the superposition of compact components, each of which has a spectrum which is cut off at long wavelengths. In general, the more compact a component is, the shorter the wavelength of its cutoff. As one observes at progressively shorter radio wavelengths, the appearance (both spectral and structural) of a quasar is essentially the same, except that it becomes smaller in angular size. One is tempted to use the common analogy of an onion skin; the major question is: at what wavelength has one 'peeled' off enough layers to finally view the radio core of the quasar? Since the flux density of a radio-loud quasar is much less in the optical than in the radio, it is known that somewhere between the radio and IR the

entire source becomes transparent, and therefore the core should be visible. However, only observations at millimetre and submillimetre wavelengths are capable of finding the longest wavelength at which the core is visible; this parameter is essential in order to determine the physical properties of the core.

Another property of many radio-loud quasars is the rapid variability of their radio and optical emission. The changes in brightness have so far defied theoretical analysis. Furthermore, the variations seen in the optical do not, except in a few cases, seem to correspond to those found in the radio. Again, the intermediate millimetre and submillimetre wavelengths are most likely to provide the key to these puzzling observations.

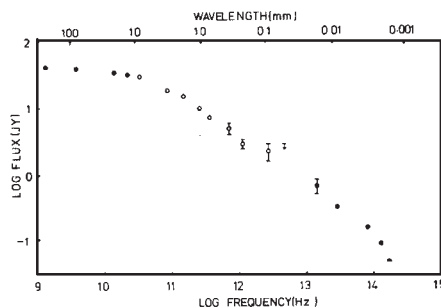


Fig. 1. The radio to infrared spectrum of the quasar 3C273. From Clegg *et al.* (*Astrophys. J.*, in the press).

While the best studied quasars are usually strong radio sources, most quasars are 'radio-quiet' (and more properly called 'quasi-stellar objects' or 'QSOs'). Searches for radio emission from objects thought to be radio-quiet QSOs show that most of them are, indeed, so weak in the radio that even the most sensitive radio telescopes cannot detect emission. Yet many of these QSOs are relatively bright in the optical. Why they should be radio-quiet remains a mystery. Determination of the wavelength beyond which the emission cuts off, apparently at submillimetre wavelengths, will help provide an answer.

The first quasar whose entire spectrum from radio to optical has been measured is the bright object 3C273. Millimetre and submillimetre observations are appearing in a pioneering paper from P. Clegg *et al.* (*Astrophys. J.*, in the press). They used the 36-foot dish of the US National Radio Astronomy Observatory at Kitt Peak, at

Alan P. Marscher is an Assistant Professor in the Department of Astronomy at Boston University, Massachusetts 02215.

millimetre wavelengths, the 3.8-metre United Kingdom Infrared Telescope (UKIRT) atop Mauna Kea, at sub-millimetre wavelengths, and NASA's Kuiper Airborne Observatory, at far-IR wavelengths.

The resultant radio to optical spectrum (see the figure) shows that the quasar has a flat radio spectrum and becomes transparent at millimetre and shorter wavelengths. It therefore appears that at millimetre and submillimetre wavelengths we are observing the core of the radio emission. This observation makes it all the more worthwhile to develop millimetre-wavelength very-long-baseline interferometry (VLBI), which will allow the angular structure of the core to be determined.

Studies of the variability of quasars at millimetre wavelengths have been rather few until recently. Now, two groups, E. Epstein *et al.* (*Astr. J.* **87**, 499; 1982) and Ennis, Neugebauer and Werner (*Astrophys. J.* **262**, 451; 1982), have published monitoring projects at 3.3 and 1 mm respectively. Unfortunately, new observations only serve further to subvert straightforward interpretation of the observed brightness changes. Sometimes the millimetre variations are correlated with those seen at longer wavelengths, and sometimes they are not. Even when they are connected, the simple models found in the literature do not seem to explain the details of the observations. Moreover, the connection with optical brightness changes is far from clear. Clearly, further observational and theoretical advances are necessary.

Neugebauer and Werner (*Astrophys. J.* **262**, 460; 1982) bears on the radio-quiet QSO problem. They detected no radio-quiet QSOs at 1 mm, so clearly the optical-IR spectrum must cut off at wavelengths substantially shorter than this.

A further curious result is the failure to detect QSOs that have been observed to be weak radio sources, implying that the radio spectrum does not rise steeply from long to short wavelengths as would be the case for a self-absorption cutoff. This behaviour is consistent with Scheuer and Readhead's (*Nature* **277**, 182; 1979) relativistic beaming model, but other observations seem to indicate that this is not the entire explanation.

Fortunately, we can expect further development of millimetre and sub-millimetre astronomy in the near future. Resurfacing and expansion of the NRAO 36-foot dish (now called the '12-metre' telescope) and the introduction of new millimetre-wave dishes (for example, the upcoming British 15-m dish atop Mauna Kea) should at least partially compensate for the demise of the proposed US 25-m project. Continued advances in sub-millimetre detectors and the deployment of IR satellites will further assure a place for millimetre and submillimetre astronomy at the forefront of quasar research. □

Vaccine development

Hybrid vaccinia virus for mass hepatitis immunization?

from A.J. Beale

IN this issue of *Nature* (p.490), Smith, Mackett and Moss from the National Institutes of Allergy and Infectious Diseases in Bethesda, Maryland, describe the successful selection of vaccinia virus into which the gene for hepatitis B virus surface antigen has been inserted. The modified virus grows well in cell culture and hepatitis B virus agglutinin (HBsAg) is released into the culture fluid in a form indistinguishable from the native antigen present in the serum of carrier individuals. Even more exciting is their observation that rabbits infected with the hybrid vaccinia virus produce high titres of antibody to the HBsAg, more than those required to provide protection in man. The authors suggest that this type of hybrid vaccinia virus might be ideally suited to mass immunization campaigns in areas of the world where control measures are urgently needed but existing vaccines are too expensive to make or administer. The success of the WHO campaign for the global eradication of smallpox gives added force and credibility to that point of view.

Moss and his colleagues devised a means of introducing foreign DNA into vaccinia virus so that the polypeptide encoded by the new DNA is produced during the growth cycle of the virus. A chimaeric gene consisting of vaccinia virus promoter sequences fused to the coding sequence for the desired foreign protein was flanked by vaccinia virus DNA in a plasmid vector. One such foreign protein used in early experiments was herpes virus thymidine kinase. In order to make vaccinia useful as a general expression vector, plasmids with restriction enzyme sites next to the vaccinia

thymidine kinase promoter or another vaccinia promoter engineered in the thymidine kinase gene were constructed. Transfection of vaccinia virus-infected cells with the plasmid allowed recombination to take place. The recombinants containing the foreign protein-coding sequences could be selected by looking for thymidine kinase vaccinia plaques or the presence of herpes virus thymidine kinase.

The recombined vaccinia containing the HBsAg-coding sequences were used to vaccinate two rabbits; typical vaccinia lesions were produced and, in addition, high titres of antibodies were seen to last for at least 31 days. These preliminary results hold promise that many other antigens may be expressed in the same way and used as vaccines, either as a means of producing antigens in culture or, more imaginatively, to use them as combination vaccines. Several problems in developing vaccines by new technology may thus be solved. Although rDNA techniques should enable protective antigens to be produced in large quantities, some antigens are poorly expressed or poorly immunogenic when produced in *Escherichia coli*, as is the case with foot-and-mouth disease virus VP1 or HBsAg. Again, the presentation of the immunogen to the host may be below the optimum, and it is commonly thought that a new generation of adjuvants or a deeper understanding of the nature of immunogenicity will be required before it will be easy to elicit the range of immune responses necessary for clinical immunity.

The cost of vaccines, and particularly of their administration, are vital factors in determining the success of an immuniza-

Oncogenic Intelligence

Oncogene, not polymorphism

from Peter Newmark

THE 15 April issue of *Science* contains a retraction of a paper published there two months ago and the subject of this column in *Nature* **301**, 654 (1983). In their paper (*Science* **219**, 853; 1983) R.J. Muschel, G. Khoury, P. Lebowitz, R. Koller and R. Dhar claimed that polymorphism rather than somatic mutation accounted for the single base change that distinguishes the *c-ras*^H oncogene of T24 bladder carcinoma cells from that in normal tissue gene banks and accounts for the ability of the former to transform NIH 3T3 cells.

At best their evidence — for the same base change in the *c-ras*^H gene of both normal and tumour tissues of an individual

with bladder cancer as in T24 cells — was decidedly preliminary. At worst, as they have since discovered, it was all accounted for by contamination. Further tests have convinced them that neither normal nor tumour tissue of their patient had the mutation that characterizes the *c-ras*^H oncogene of T24 cells.

Polymorphism remains a possible explanation of the T24 mutation but the attempt by Muschel *et al.* to make that explanation stick should now be counted as a casualty of the fierce competition that abounds between oncogene laboratories. □

Peter Newmark is Deputy Editor of *Nature*.

tion campaign, especially in developing countries. The use of vaccinia as a vehicle might solve both these problems. Hepatitis B vaccine derived from blood is prohibitively expensive but the cost of vaccinia-expressing HBsAg should be little more than regular vaccinia virus; and, moreover, the methods of administration pioneered by the WHO smallpox eradication programme are very cost effective. The bifurcated needle developed by Wyeth Laboratories is easy to use by relatively unqualified vaccinators and the vaccination scar is a clear sign of a previous successful vaccination.

Some difficult problems remain to be solved before this strategy can be adopted. In order to eradicate smallpox, a high proportion of the world population was vaccinated with vaccinia virus and will, if revaccinated, have immunity or partial immunity (Fenner, *F. Prog. med. Virol.* 23, 1; 1977) which will limit the growth of vaccinia virus and the production and response to HBsAg. The scope of any new campaign based on vaccinia virus will thus be reduced. The immunizing effect of the vaccinia virus vehicle against itself might also limit its usefulness as a general method; even supposing it could be used

for hepatitis B immunization for a defined population group, it would not be available for use with another immunogen. To overcome this problem, a series of antigenically distinct, safe attenuated viruses would be required, but that would be by no means easy to supply.

A second major problem is safety. In several Western countries the danger from vaccinia virus immunization was thought too great relative to the risk from smallpox at the time the WHO eradication campaign was launched (Dick, G. *Prog. med. Virol.* 8, 1; 1966). In the USA and UK, for example, the risk of death from imported smallpox was already extremely low. The great advances in molecular genetics that are now taking place, however, should make it possible to manipulate the virus to avoid these drawbacks, or to produce a series of attenuated viruses colonizing the mucous membranes of the intestinal or respiratory tract. Adenovirus, for example, might be used to induce local immunity at the portal of entry as well as circulating antibody and cell-mediated immunity. □

A. J. Beale is Director of Biological Products at the Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS.

Latent herpes simplex virus

The state of the herpes genome

In this issue of *Nature* (p.523), Daniel Rock and Nigel Fraser of the Wistar Institute report the successful detection of herpes simplex virus (HSV) DNA in the latently infected central nervous tissue of mice. The new results provide the first step in understanding the nature of the latent virus; it seems that although the whole of the herpes virus genome is present in the latent form, it is qualitatively different from that in virus particles or infected cells.

Since the introduction of animal models to the study of HSV latency and the demonstration that sensory ganglia are the prime source of latent virus¹ the question of the nature of the latent virus has been a fascinating but frustrating one. The assay for latent virus, the detection of infectious virus following culture of explanted neural tissue but not by direct homogenization and assay, implies the presence of the virus genome in a non-infectious form during latency. It can be argued, however, that explant culture is merely a more sensitive assay, allowing the detection of very small numbers of infectious particles resulting from a low-level chronic infection; support for this view comes from the electron microscopic observation of virus particles in the trigeminal ganglia of latently infected rabbits².

Evidence against a chronic infection or

'dynamic' latency comes from three types of experiment. First, latent infection has been established using temperature-sensitive mutants in tissues whose temperatures are non-permissive for mutant replication^{3,4}. Second, antiviral drugs that efficiently suppress virus replication during the acute phase of infection fail to 'cure' animals of established latent infection⁵⁻⁷. Third, HSV DNA was detected in latently infected tissue by kinetic hybridization⁸, but no virus-specific RNA was found, a result in contrast to that obtained with tissue taken from the acute phase of infection. These observations argue against viral replication being a requirement for maintenance of the latent infection and favour a 'static' model.

Attempts to examine the form of HSV DNA in latently infected tissue have been bedevilled by the small amount of virus DNA present and by the considerable cross-hybridization of HSV DNA sequences with mammalian cell DNA⁹⁻¹¹. Success has finally come in the new work reported by Rock and Fraser. Virus sequences were detected in DNA from the central nervous system and trigeminal ganglia of mice several months after corneal inoculation of HSV-1. Using virus DNA as a hybridization probe they were able to show that the entire virus genome was present. Of particular interest is that when a DNA fragment composed of the internal repeated sequences was used to

probe blots of restricted DNA, the probe detected both internal and terminal repeats in DNA from the acute phase of infection but only internal repeats in DNA from the latent phase. The virus DNA in latently infected mice thus appears to be 'endless'. The data are of course open to several interpretations since 'endless' DNA could result from the formation of circles or long concatamers or from the random integration of linear molecules. Nevertheless, these are the first data that demonstrate that 'latent' HSV DNA molecules are qualitatively different from the molecules found in virus particles or in infected cells.

Rock and Fraser have combined sensitive hybridization techniques with an animal model in which relatively large amounts of viral DNA persist in the latent phase of infection, so further data should soon be forthcoming. It will be particularly interesting to see whether the viral DNA in latently infected animals is present in more than one of the four isomeric forms found in virus particles.

A note of caution should perhaps be sounded about the use of the term 'latent virus DNA'. Much of Rock and Fraser's evidence derives from experiments with DNA from the brain stem of latently infected animals and, while there is considerable evidence that the central nervous system can act as a site of latent HSV, *in vitro* reactivation of virus from the central nervous system is very inefficient by comparison with reactivation from the sensory ganglia¹². Indeed, Rock and Fraser were unable to reactivate virus from the brain stem of latently infected mice despite the presence of detectable HSV DNA, a result previously reported by Cabrera *et al.*¹³. Thus while the brain stem is a valuable source of material for biochemical studies, it lacks authenticity as a 'latently infected tissue'.

Finally, two lines of evidence suggest that limited viral gene expression occurs in latently infected cells. Galloway *et al.*¹⁴ have detected virus-specific transcripts in neurones from human sensory ganglia by *in situ* hybridization, while Green *et al.*¹⁵ have detected an immediate-early polypeptide of HSV in the sensory neurones of latently infected rabbits by immunofluorescence. Unfortunately, the transcripts detected by Galloway *et al.* are not derived from that part of the HSV genome that encodes the immediate-early polypeptide in question. This problem, like so many in this field, remains to be resolved. □

A. C. Minson is in the Department of Pathology, University of Cambridge, Hills Road, Cambridge CB2 2QQ.

1. Stevens & Cook *Science* 173, 843 (1971).
2. Baringer & Swaveland *Lab. Invest.* 30, 230 (1974).
3. Lofgren *et al. Virology* 76, 440 (1977).
4. McLennan & Darby *J. gen. Virol.* 51, 223 (1980).
5. Field *et al. Antimicrob. Ag. Chemother.* 15, 554 (1979).
6. Klein *et al. Antimicrob. Ag. Chemother.* 15, 723 (1979).
7. Blyth *et al. J. gen. Virol.* 48, 417 (1980).
8. Puga *et al. Virology* 89, 102 (1978).
9. Maitland *et al. J. gen. Virol.* 55, 123 (1951).
10. Peden *et al. Cell* 31, 71 (1982).
11. Puga *et al. Cell* 31, 81 (1982).
12. Cook & Stevens *J. gen. Virol.* 31, 75 (1976).
13. Cabrera *et al. Nature* 288, 288 (1980).
14. Galloway *et al. J. Virol.* 41, 686 (1982).
15. Green *et al. Infect. Immun.* 34, 987 (1981).

Particle physics

After the W boson . . . the Z?

from Frank Close

THIS week physicists at CERN, Geneva hope to begin the second round of experiments that could prove definitive in isolating the W and Z bosons — the carriers of the weak nuclear force. The results of the first experiments were announced in January; the media announced the discovery of the W boson, two independent groups of theorists claimed that the top quark had also been seen, but the experimentalists (who are, of course, closest to the data) have so far only claimed to find "electrons with large transverse energy and missing associated energy . . . which have the signature of a two body decay . . . fitting well the hypothesis that they are produced by a W^+ or W^- boson". The interpretation is that the missing energy is carried by an undetected neutrino, the W having decayed into the electron and the neutrino.

The events are singularly impressive, yet, as Norman Dombey (*New Scientist* 17 February 1983) reminds us, "the experimentalists do not see a W *per se* . . . a single electron cannot identify a W with certainty". This caution, correctly exercised in the two experimental papers (UA1 and UA2 collaborations *Phys. Lett.* 122B, pages 103 and 476 respectively) has been relaxed during recent weeks as more and more physicists received detailed first-hand seminar reports about the discovery. There is no doubt that the two teams have seen a new phenomenon and no alternative mechanisms can satisfactorily explain them.

The abundance of the events and also their electron energy distribution are as expected if a W boson has been produced and has a mass of 81 ± 5 GeV (UA1 report) or 80 ± 10 GeV (UA2 report). This is truly exciting, for the unification of weak and electromagnetic forces can be achieved if the W mass is 82 ± 2 GeV.

Nuclear γ decay is an 'electromagnetic' process; β decay, by contrast, is brought about by the weak force. The probability of neutron β decay into a proton, an electron and an antineutrino is controlled by Fermi's weak coupling constant G ; electromagnetic effects, by contrast, involve the fine-structure constant α . The weak force, as its name implies, appears to be more feeble than the electromagnetic force and its range is much less. In β -decay and other low-energy phenomena the two forces appear to be unrelated.

A quantitative measure of their relative apparent strengths is given by the ratio of $G/\alpha \approx 10^{-3}$ at an energy of 1 GeV. This rather opaque caveat 'at 1 GeV' is important: G has dimensions of (energy) $^{-2}$ whereas α is dimensionless $\approx 1/137$. The dimensionless quantity $G \times (\text{energy})^2$ can be sizeable at high energy and comparable with α in magnitude.

It was realised that if effects due to the absence of mirror symmetry in weak interactions were allowed for, then electromagnetic and weak interactions could have identical intrinsic strength — provided the weak force contains an important 37

GeV energy scale. In low-energy experiments, such as neutron β -decay, this 37 GeV is hidden and the force appears to be weak ($\approx 10^{-5}$). But at high energies the 37 GeV will be out in the open and the true strength α will be revealed.

The 37 GeV is a measure of the mass of the weak force quantum (the analogue of electromagnetism's massless photon) — the W boson. In β decay, energy-momentum conservation prevents creation of a 37 GeV quantum — the uncertainty principle can, however, allow it to exist fleetingly. Indeed, its existence would be so Pickwickian at low energies that it can transmit the force for only 10^{-15} cm. The feeble strength and limited range of the weak force at low energies can thus both be understood.

The germ of this theory was put forward by Julian Schwinger in 1957 and contained γ and W^+ , W^- . Two decades of theoretical work on gauge field theories (of which quantum electrodynamics is the simplest example) and an increased awareness of the role of hidden symmetries (spontaneous symmetry breaking) in renormalizable field theories, culminated in the powerful modern theory due to Sheldon Glashow, Abdus Salam and Steven Weinberg.

The theory requires γ , W^+ , W^- and also a Z^0 boson, the latter being the carrier of a new weak force (discovered in 1973). Clebsch Gordan coefficients and other numerical factors cause the old 37 GeV to be modified as follows. The weak forces have intrinsic electromagnetic strength if

$$W^+ = W^- = 82 \pm 2 \text{ GeV}$$

$$Z = 92 \pm 2 \text{ GeV}$$

Other aspects of the theory relate the W and Z masses, so that if the theory is right, and if the W has indeed been detected at about 80 GeV, then the Z^0 must exist with the cited 90 GeV mass.

The good news about Z^0 is that it can decay into e^- and e^+ which, being charged, are easy to detect. Hence a Z^0 can be cleanly and unambiguously identified if it is produced in the experiment. The bad news is that this type of event is expected to be only one-tenth as likely as the W events — as only nine of the latter have been seen so far, it is no surprise that the Z^0 is still undetected.

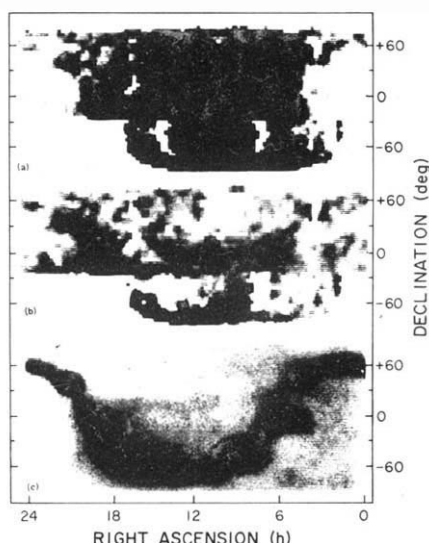
But this should change when the new experiments are started. With improved efficiency of the accelerator the optimists hope that up to 100 W and 10 Z^0 events will be found in the modes $W^+ \rightarrow e^+ + \text{neutrino}$; $Z^0 \rightarrow e^+ e^-$. The masses, production rate, lifetime of W and Z and the angular distributions of the decay electrons are all predicted. The $Z^0 \rightarrow e^+ e^-$ events provide a potentially clean unambiguous signature. A larger sample of W candidates will also generate more confidence in their interpretation. Not least, the ratio of W and Z masses is a crucial prediction of the model. It may be possible to spot other decay modes of the W.

In addition to the events containing an isolated electron moving transverse to the beams with no visible energy balance (the

Cosmology

Quadrupole fades away

THERE is, after all, no quadrupole component currently detectable in the cosmic microwave background, according to two recent papers^{1,2}. The result is significant for cosmologists who, following earlier reports of a possible quadrupole, were relating such inhomogeneities to distributions of matter in the early Universe³. Analysis of the data shown here² failed to confirm the reported quadrupole, which is now attributed to experimental artefacts. (a) A map of the sky at 24.5 GHz showing the dipole effect caused by the Earth's motion relative to the 2.7K microwave background. The measurements were made during three flights of a balloon-borne radiometer. The gaps in the data result from malfunction and the limited sky coverage available. (b) The same data with a best-fit dipole subtracted, revealing the galactic plane. (c) A map derived from a model for galactic emission at this frequency.



Phillip Campbell, *Physical Sciences Editor of Nature*.

1. Lubin, P.M., Epstein, G.I. & Smoot, G.F. *Phys. Rev. Lett.* 50, 616 (1983).
2. Fixsen, D.J., Cheng, E.S. & Wilkinson, D.T. *Phys. Rev. Lett.* 50, 620 (1983).
3. Muller, R.A. *Nature* 291, 609 (1981).

W candidate events), the experimenters also reported a dozen events where a particle, probably an electron or, positron, emerges opposite a jet of pions and kaons. If all these events do contain electrons or positrons, then they have the character expected of top quarks undergoing β decay. The possibility that particles containing the elusive top quark are being produced would be an exciting, if not unexpected, bonus.

The familiar β decay involves the first generation of quarks: down $\rightarrow$ up + ($e^- \bar{\nu}$) (*Nature* 271, 406; 1978). The β decay of charmed particles involves the second generation: charm $\rightarrow$ strange + ($e^+ \nu$). If a third generation of quarks (top and bottom) exists, then its β decay will be top $\rightarrow$ bottom + ($e^+ \nu$). Have such events been seen?

The proton-antiproton collisions at CERN could occasionally produce a top particle which β -decays, producing an energetic positron on one side of the beams and a jet of particles from the subsequent bottom decay on the other side. This event was widely predicted before the experiments. Theorists in Britain, US and India have independently noticed that the 'electron plus jet' events, suitably interpreted, could indeed be indicative of the top quark's existence with a mass of between 25 and 40 GeV.

The best place to produce top quarks would be in electron positron annihilations, but no machine exists that has enough energy to produce such massive objects (though the 25 GeV possibility is tantalizingly near to the upper energy of the PETRA facility in Hamburg). At present only the CERN facility and future $e^- e^+$ machines could do so.

But it is moot whether or not the theorists have jumped the gun. The events that they used to claim evidence for top quarks were published in the W paper and thus are freely available. However, they are more complicated than the W events and were only analysed to the level required for them to be used as a background against which the candidate events could be contrasted; it might be premature to use them as demonstrating specific new phenomena. The experimentalists are now studying them in more detail but have not yet published their conclusions.

At the very least the theorists have sharpened some of the issues at stake in the search for top quarks: subsequent data might even confirm their proposal. During the next three months, further examples of the 'electron plus jet' events will be looked for as intently as the isolated electron (W) and $e^- e^+$ pair (Z). By July 1983 we might have witnessed as many milestones as in the *annus mirabilis* 1932. Watch this space. $\square$

Frank Close is Senior Principal Scientific Officer at the Science and Engineering Research Council Rutherford Appleton Laboratory, Chilton, Oxfordshire OX11 0QX.

Population genetics

Adaptive evolution in the laboratory

from Brian Charlesworth

A PREDICTION of the theory of evolution by natural selection is that most mutations should be harmful, since contemporary species are the product of a long history of evolutionary change in which most useful mutations would by now have been fixed by selection¹. There is overwhelming evidence that this is in fact the case, especially for *Drosophila*, where a large research effort has been devoted to measuring the effects of new mutations on fitness². The challenge of a new environment should, however, enable some previously harmful mutations to acquire a selective advantage, and the familiar cases of rapid evolutionary change involving single genes, such as industrial melanism³ and pesticide resistance⁴, are examples of this. It is likely, however, that many subtle changes, with no obvious effects at the gross phenotypic level, take place when a population adapts to a new environment, but these usually go undetected.

A method of detecting such changes is provided by the phenomenon known to microbial geneticists as periodic selection^{5,6}. So far, periodic selection has only been studied in prokaryotes, but in a paper published in this issue of *Nature* (p.495), Charlotte Paquin and Julian Adams of the University of Michigan⁷ describe experiments on a eukaryote, the yeast *Saccharomyces cerevisiae*. Studying periodic selection in yeast made it possible for them to perform an important experiment — a comparison of the rates of evolution in haploid and diploid asexual populations.

Periodic selection may be understood as follows. Consider an asexually reproducing population of a microorganism, maintained in a chemostat (a device that provides a continuous flow of nutrients into a population of constant size). The experimenter ensures that the population contains a low frequency (say one in a million) of an easily scored mutant gene at a certain locus, which is approximately selectively neutral in the chemostat conditions, for example, a gene conferring resistance to a poison absent from the culture medium. The frequency of this marker gene will increase linearly over time, at a rate equal to the rate of mutation of the wild-type allele to the marker gene. If a selectively favourable mutation occurs at some other locus, it is almost certain to appear in a cell lacking the marker gene, since the gene is extremely rare. The asexual mode of reproduction means that the advantageous mutation is completely associated with the wild-type allele of the marker gene, so that it causes a decline in

frequency of the marker when it starts to increase in frequency under selection. Once the clone containing the favourable mutation has risen to a high frequency, it will accumulate mutations to the marker gene, which will therefore reverse the decline in frequency. A sequence of events involving the spread of favourable mutations thus induces a zig-zag curve of frequency of the marker gene over time (see Fig.1, on page 496), each downward limb of the curve indicating that an advantageous mutation has succeeded in establishing itself at a significant frequency in the population.

In the experiments described by Paquin and Adams in this issue of *Nature*, rates of adaptation to a new environment, a carbon-limited chemostat, were compared for diploid and haploid yeast populations. An important question that can be answered using this comparison concerns the level of dominance of advantageous mutations: in very large diploid populations (as in these experiments, which involve populations of the order of 10^9 cells), the chance of establishment of a completely recessive advantageous mutation is effectively zero⁸, since individuals expressing the effects of the gene are produced only by double mutation events. Adaptive evolution would thus be impossible if it relied on recessive mutations, in these conditions. (It is, of course, well known that the mutations involved in the examples of rapid evolution cited above tend to express their favourable effects when heterozygous^{3,4}.) The problem of calculating the expected rate of evolution in an asexual population is a tricky one. It is not necessarily the case that rates are directly proportional to the average selective advantage of a new mutation, or to the overall rate of mutation to favourable alleles^{9,10}, and this complicates the quantitative interpretation of the results.

Paquin and Adams observed in total the spread of 64 mutations over 2,612 generations, corresponding to a rate of spread of 5.68×10^{-12} mutations per cell per generation for diploid populations, and 3.55×10^{-12} for haploid ones. Diploids thus seem to evolve about 1.6 times as fast as haploids, on this scale. This is presumably due to their having twice as many genes available for mutation as haploids, coupled with nearly complete dominance of the favourable mutations. The conclusion about dominance was tested by doing competition experiments between strains

Brian Charlesworth is in the School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG.

carrying different numbers of advantageous mutations, as estimated from the periodic selection experiments. These experiments showed that the average selective advantage of a single favourable mutation was 0.10 for the haploid and 0.09 for the diploids (not significantly different), again suggesting a high level of dominance for the favourable mutations.

These results suggest that diploid populations may evolve faster than haploid ones, at least under circumstances where mutation rates limit the rate of evolution, and Paquin and Adams suggest that this provides an evolutionary advantage for diploidy. This is, of course, a long-term advantage at the level of the population and does not entail an immediate selective advantage to diploid variants arising in a haploid population. (Another difficulty is that the diploid cells were larger in size and smaller in number than the haploid ones, so that the difference in rate almost disappears when expressed on the scale of numbers of mutations per population.) Another important feature of these results

is that the rates of occurrence of favourable mutations are sufficiently high that only a short interval (roughly 40 generations, on average) separates successive spreading events. This is relevant to theories of the evolutionary advantages of sexual reproduction and genetic recombination, since recombination can have no influence on the rate of evolution if each successive favourable mutation becomes fixed in the population before the next one appears^{10,11}. □

1. Muller, H.J. in *The New Systematics* (ed. Huxley, J.S.) 185 (Systematics Association, London, 1940).
2. Simmons, M.J. & Crow, J.F. *A. Rev. Genet.* **11**, 49 (1977).
3. Kettlewell, H.B.D. *The Evolution of Melanism* (Oxford University Press, 1973).
4. Brown, A.W.A. & Pal, R. *Insecticide Resistance in Arthropods* (WHO, Geneva, 1971).
5. Atwood, K., Schneider, L. & Ryan, F.J. *Cold Spring Harb. Symp. quant. Biol.* **16**, 345 (1951).
6. Novick, A. & Szelard, L. *Proc. natn. Acad. Sci. U.S.A.* **36**, 708 (1950).
7. Paquin, C. & Adams, J. *Nature* **302**, 495 (1983).
8. Haldane, J.B.S. *Proc. Camb. phil. Soc.* **28**, 838 (1927).
9. Kimura, M. & Ohta, T. *Theoretical Aspects of Population Genetics* (Princeton University Press, 1971).
10. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
11. Fisher, R.A. *The Genetical Theory of Natural Selection* (Oxford University Press, 1930).

Mimicry

Beetles disguised as ants

from Robert M. May

THE availability of food within colonies of ants, termites and other social insects has led many other arthropod species to evolve ways of life that are essentially parasitic on such colonies. Wilson¹ has given a tabular summary of the 17 orders, 120 families and hundreds of genera of arthropods that live as symbionts in colonies of social insects, and he emphasizes that this list continues to grow at such a rate as to suggest that only a small fraction of all social parasites are currently catalogued. For ant colonies, such 'myrmecophiles' (ant lovers) have adopted a wide variety of devices in order to pass unnoticed among their hosts; the devices are, as it were, concrete realizations of the powers for invisibility or disguise conferred by the Niebelung's legendary Tarnhelm.

Many species of beetle, for example, exhibit morphological mimicry, with slender body forms and colouration that match their hosts. The emphasis on visual similarity is most common among parasitic species that spend a lot of time with their hosts in the open (as do many beetles that march with army ants). Parasitic species that live among their hosts underground tend to rely more upon breaking the chemical or behavioural codes of communication within the host species than upon morphological similarity. Wilson¹ summarizes instances where larval or adult stages of myrmecophilic species produce pheromones or other chemicals that elicit

particular behaviour — often regurgitation and feeding — from their duped hosts.

In general, the odour of a colony of social insects has both innate and acquired components². Thus ants appear to recognize each other by one individual touching its antennae to another's cuticle, which suggests the cuticle is a source of innate species-specific chemicals. The relatively large surface of the cuticle can also serve to absorb the acquired component of the colony odour from the surrounding environment. In particular, the cuticles of dipteran, lepidopteran and hymenopteran species have been shown to contain hydrocarbons that influence species-specific behaviour in various ways. It has been suggested that such cuticular hydrocarbons function as cues for caste and species recognition in many termite colonies³. One way of moving unremarked among one's hosts may thus be to synthesize the appropriate cuticular hydrocarbons. An example is provided by the beetle *Trichopsenius frosti*, which is a parasite in colonies of the termite *Reticulitermes flavipes*; the beetle synthesizes cuticular hydrocarbons that are identical to those of its hosts⁴.

Vander Meer and Wojcik⁵ have recently published a detailed study of the part played by cuticular hydrocarbons in the coexistence of a myrmecophilic beetle, *Myrmecaphodius excavaticollis*, with any of four species of ants in the genus *Solenopsis* (*S. richteri*, *S. invicta*, *S. geminata*, *S. xyloni*). The study of colony odours was largely motivated by the obser-

vation that the beetle is accepted into the various ant colonies despite its lack of morphological mimicry or chemical defences. Adult beetles move freely among the host ants, getting "food directly from workers through trophallaxis (exchange of alimentary liquid among colony members, either reciprocally or unilaterally), by predation on ant larvae, and by feeding on freshly killed or decomposed workers and ant booty"

Using gas chromatography, Vander Meer and Wojcik show that the cuticles of individual beetles taken from laboratory colonies of *Solenopsis* contain the hydrocarbons specific to the particular ant species: each of the four *Solenopsis* species has its own distinct set of cuticular hydrocarbons and its own distinct gas chromatographic signature. Separation of *M. excavaticollis* from any ant hosts results (after about two weeks) in loss of most of the host hydrocarbons, leaving a pattern of cuticular hydrocarbons that is innate to the beetle and quite distinct from any one of the four ant species. Transfer of a beetle from one ant colony to another of a different species results in its acquiring the cuticular hydrocarbons of the new host: typically, 15 per cent of the beetles' cuticular hydrocarbons are found to be host hydrocarbons within two hours of its introduction into an ant colony; the figure rises steadily to around 50 per cent within three to four days. During the first few days following introduction into an ant colony, *M. excavaticollis* individuals are attacked, and they respond by 'playing dead' and relying on their armoured exterior to protect them.

In view of the diversity of cuticular hydrocarbon patterns that *M. excavaticollis* showed itself capable of acquiring, and the manner in which such patterns built up following introduction of the beetle into a new ant colony, Vander Meer and Wojcik conjectured that the hydrocarbons are gained passively from ant-beetle contact, rather than being actively synthesized by the beetle. The idea was confirmed by placing freshly killed *M. excavaticollis* into colonies of *S. invicta*: the corpses acquired the characteristic hydrocarbon signature of the ant colony within two days.

Vander Meer and Wojcik's study adds one more strategy — which they call 'passive integration' — to the array of tricks that myrmecophiles use. Such passive acquisition of the host colony's odour, coupled with the beetles' armoured exterior, is an effective method for coping with a diverse assortment of host species. □

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, New Jersey 08544.

1. Wilson, E.O. *The Insect Societies* Ch. 20 (Harvard University Press, 1971).
2. Jutsunt, A.R., Saunders, T.S. & Cherrett, J.M. *Anim. Behav.* **27**, 839 (1979).
3. Howard, R.W. & Blomquist, G.J. *A. Rev. Ent.* **27**, 149 (1982).
4. Howard, R.W., McDaniel, C.A. & Blomquist, G.J. *Science* **210**, 431 (1980).
5. Vander Meer, R.K. & Wojcik, D.P. *Science* **218**, 806 (1982).

Biological applications of picosecond spectroscopy

E. F. Hilinski & P. M. Rentzepis

Bell Laboratories, Murray Hill, New Jersey 07974, USA

Technological advances in picosecond spectroscopy have permitted the mechanisms of various chemical, physical and biological processes to be elucidated and understood to a greater degree than ever before. By means of picosecond emission, absorption and Raman spectroscopy, one can probe and measure directly the transient intermediates and kinetics of primary events in complex biological processes. A description of two current types of laser systems—solid-state and synchronously pumped dye lasers—and their application to determining the primary events in the biological processes of dissociation of oxy- and carboxymyoglobin, excited-state relaxation of porphyrins and visual transduction, illustrate the power of picosecond spectroscopy.

PRIMARY events of many biological processes including photosynthesis, visual transduction and oxymyoglobin (MbO) dissociation, are known to occur within picoseconds (10^{-12} s). Spectroscopists have taken advantage of the extremely short pulses of light which emerge from mode-locked lasers to probe the dynamics of these and other rapid physical, chemical and biological processes. Various laser systems with temporal resolution in the range of less than one to several picoseconds can now be used to observe transient molecular states and species and to measure their formation and decay kinetics by means of absorption, emission and Raman spectroscopy. The elucidation of detailed mechanisms of ultrafast complex events is in principle possible by judicious utilization of these techniques. Picosecond studies have grown in number so rapidly in the past decade that we shall not attempt to discuss this vast area in general but rather present data and discuss picosecond results which pertain to the initial events leading to visual transduction, to the relaxation mechanism in porphyrins, and to the dissociation of oxy- and carboxymyoglobin. Our aim is to present some examples of the types of picosecond spectroscopy used for the study of transient species and to demonstrate the suitability of this field to many areas of chemistry, physics and biology.

Experimental methods

Lasers which produce light pulses of picosecond duration and have been used extensively to study chemical, physical and biological processes, fall into two general categories: (1) solid-state oscillators such as $\text{Nd}^{3+}:\text{YAG}$ (yttrium-aluminium-garnet) or $\text{Nd}^{3+}:\text{glass}$ and (2) those based on synchronously pumped dye lasers. Each type has desirable features as well as shortcomings. Because the solid-state systems were the first developed (and have been discussed previously^{1,2}), they have been used in most of the picosecond work performed thus far and possibly its most critical applications. More recently, synchronously pumped dye laser systems have become increasingly popular because of their tunability and production of shorter pulses.

Because the manipulation of the picosecond pulses for spectroscopy and kinetics measurements is basically the same after pulse generation and amplification has been accomplished, regardless of pulse origin, the description of the picosecond system presented in Fig. 1 may with minor variations be used to describe both the dye and solid-state lasers. It consists of an actively mode-locked argon ion laser which pumps a ring dye laser resulting in the generation of low energy pulses with a duration of ~ 1 ps full-width at half-maximum. In this particular system, the argon ion laser is mode-locked at 123 MHz by an acousto-optic modulator driven by an ultrastable frequency generator. The average power of ~ 4 W is obtained with a pulse

duration of 150 ps at 514.5 nm. These pulses subsequently pump a ring laser usually having rhodamine 6G(R6G) for the lasing medium. The pulses are generated at a repetition rate of 246 MHz with a width of ~ 1 ps and tunable in wavelength between ~ 570 and ~ 670 nm (R6G) with an average power of ~ 100 mW. A background-free autocorrelator is used to monitor the short pulses of the dye laser³.

While this method is relatively convenient and inexpensive for measuring pulse-widths of such a short duration, it can happen that more than a single picosecond pulse is present within a time (frame envelope) of ~ 10 ps and may pass undetected by the autocorrelation method⁴, giving the appearance of a shorter pulse. These multiple pulses, as well as broadened single pulses, are frequently caused by improper adjustment of the length of the ring laser cavity. Changes in length on the order of micrometres are sufficient to reduce the quality of the pulse from its optimum shape. The most reliable means to determine and to assure that there exists only one ultrashort pulse is probably a streak camera.

Spectroscopic studies for most chemical and biological studies of interest require the use of either a pulse of rather high energy (~ 1 – 10 mJ available) after intense amplification or a train of pulses with high repetition rate and lower energy content. The selection of a single pulse is accomplished by means of the usual arrangement of a Pockels cell and crossed Glan polarizers. The horizontally polarized pulses of the ring laser are transmitted by the first polarizer and Pockels cell but are rejected by the second, crossed polarizer. However, when an ~ 8 kV voltage pulse is applied to the potassium dihydrogen phosphate (KDP) crystal of the Pockels cell, the pulse present within the crystal has the direction of its polarization rotated by 90° and is therefore transmitted by the second polarizer.

This selected single pulse may be amplified by a set of dye amplifiers consisting of a series of flowing dye cells pumped by 10-ns pulses of the second harmonic of a $\text{Nd}^{3+}:\text{YAG}$ laser with a repetition rate of ~ 10 Hz. The arrival of the selected picosecond pulse from the ring dye laser at the dye cells is timed to overlap the arrival of the 532-nm pump pulse of the $\text{Nd}^{3+}:\text{YAG}$ laser. Saturable absorbers are placed between each flowing dye cell for the removal of superradiance. This pulse, which has been amplified by a factor of $\sim 10^6$, has sufficient energy to be used for excitation either directly or converted to wavelengths, which are not emitted by the dye laser, by means of frequency doubling and/or stimulated Raman scattering.

Picosecond emission spectroscopy

In using picosecond emission spectroscopy to identify the emission spectra and lifetimes of short-lived intermediates, the picosecond pulse is directed into the sample cell and the

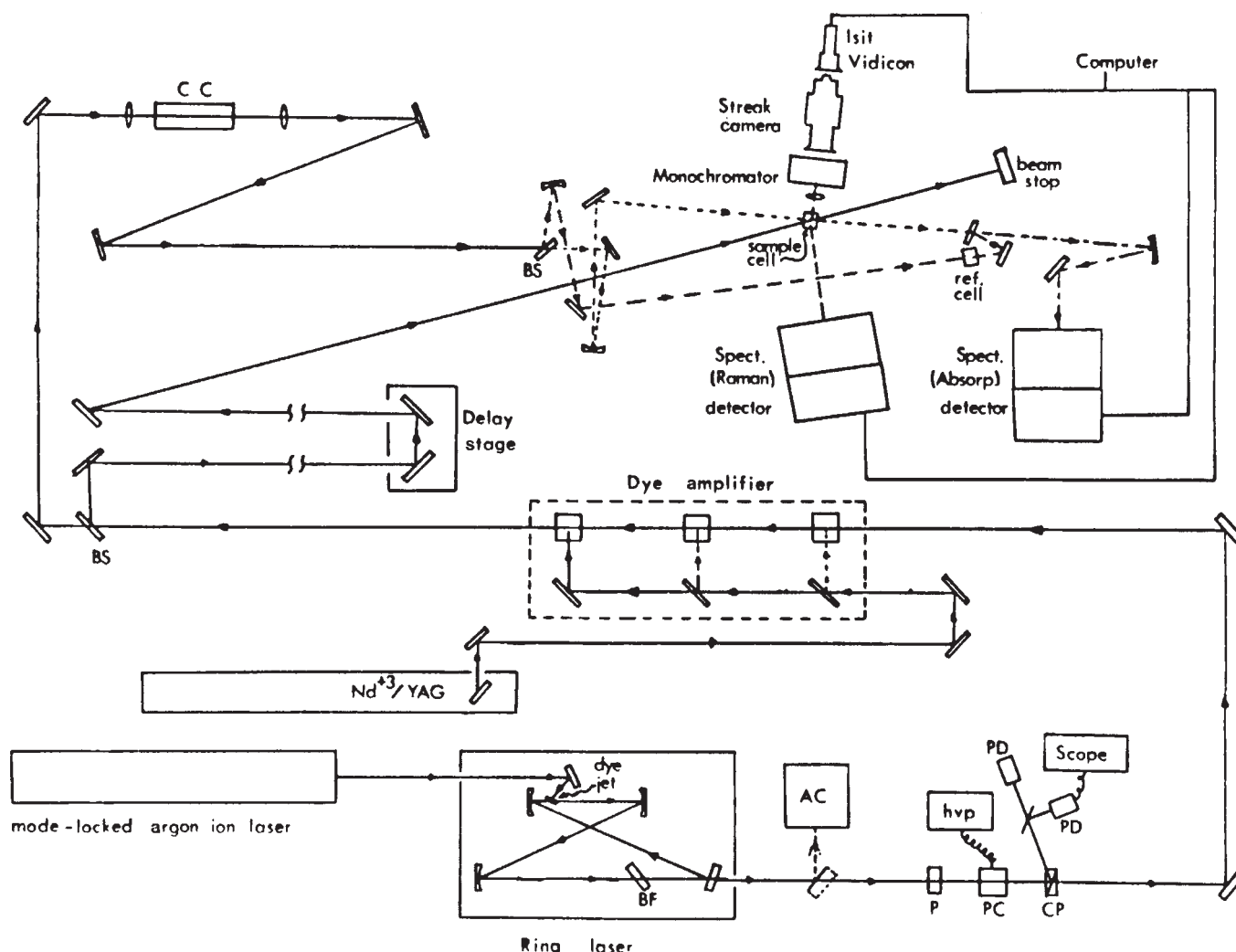


Fig. 1 Schematic diagram of a synchronously pumped dye laser system. AC, autocorrelator; P, polarizer; PC, Pockels cell; CP, crossed polarizer; hvp, high-voltage pulser; PD, photodiode; BS, beam splitter; CC, continuum cell.

emission intensity is measured temporally by means of an assembly of streak camera, vidicon, optical multichannel analyser (OMA) and minicomputer. Filters or a monochromator may be used to confine the measurement of the emission to a specific spectral region. Emission spectra of a transient molecular species may be obtained with the assembly of spectrometer, vidicon and OMA shown in Fig. 1. The data are sent from the OMA to a minicomputer for analysis, averaging, extracting kinetic lifetimes of transients and storage.

Absorption and Raman spectroscopy

To monitor a complete process, it is necessary to record, in addition to emission spectra, the transient absorption as a function of time and, whenever possible, the transient Raman spectra. For absorption and Raman spectroscopy, the optical arrangement is modified such that the pulse is split into two parts. One is used for excitation while in the second the probe pulse is used either for Raman scattering or to generate the interrogating absorption broad-band continuum. In the latter case, the probe pulse is directed through a cell containing D_2O/H_2O resulting in a broad-band continuum picosecond pulse which is consequently shaped by means of appropriate filters and used to detect absorption spectra of the transient species generated by excitation at a selected recording time after excitation. The data are detected by an assembly of spectrometer, vidicon and OMA and sent to a minicomputer for

analysis, display and storage. For Raman spectroscopy, the probe pulse arrives at a selected time after excitation, causing scattering which generates time-resolved Raman spectra. The scattered light is collected by means of an assembly similar to that used for absorption measurements and shown schematically in Fig. 1.

In the following sections we shall present a few examples of biological interest which illustrate the power of picosecond spectroscopy to identify the intermediate species involved in the course of a reaction as well as to provide kinetic information about the various evolving species.

Porphyrins and metalloporphyrins

Vibrational and electronic spectra have frequently been used to identify the structure and relaxation characteristics of porphyrins. Complexation of the cyclic tetrapyrrole ring of porphyrins greatly affects their photophysics⁵⁻¹⁰. Features of the visible absorption spectrum and the formation and decay kinetics of the excited states of porphyrins and metalloporphyrins have been understood as a result of several theoretical studies^{5,11-15}. In general, porphyrins which do not possess a metal and those which are complexed with closed-shell diamagnetic metal ions such as $Zn(II)$ and $Mg(II)$ are strongly fluorescent and possess fluorescence decays which are slower than those of other types of metalloporphyrins. Metalloporphyrins formed by complexation with open-shell diamagnetic ions such

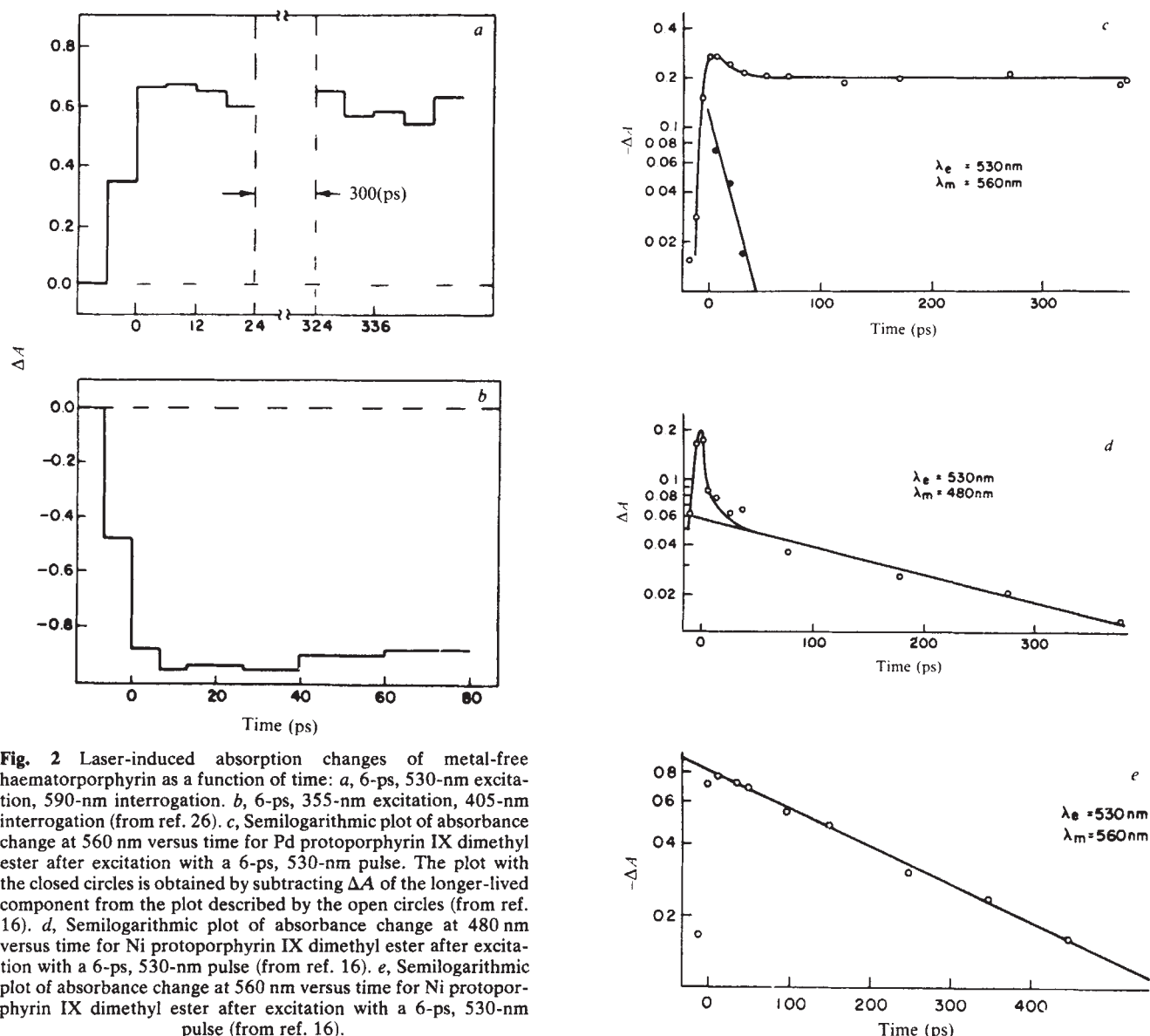


Fig. 2 Laser-induced absorption changes of metal-free haematorporphyrin as a function of time: *a*, 6-ps, 530-nm excitation, 590-nm interrogation. *b*, 6-ps, 355-nm excitation, 405-nm interrogation (from ref. 26). *c*, Semilogarithmic plot of absorbance change at 560 nm versus time for Pd protoporphyrin IX dimethyl ester after excitation with a 6-ps, 530-nm pulse. The plot with the closed circles is obtained by subtracting ΔA of the longer-lived component from the plot described by the open circles (from ref. 16). *d*, Semilogarithmic plot of absorbance change at 480 nm versus time for Ni protoporphyrin IX dimethyl ester after excitation with a 6-ps, 530-nm pulse (from ref. 16). *e*, Semilogarithmic plot of absorbance change at 560 nm versus time for Ni protoporphyrin IX dimethyl ester after excitation with a 6-ps, 530-nm pulse (from ref. 16).

as Pd(II) or Pt(II) fluoresce only slightly, if at all, but usually exhibit strong phosphorescence. Conversely, porphyrins which incorporate paramagnetic ions such as Cu(II) or Fe(II) phosphoresce only slightly, and do not fluoresce at all. Kobayashi *et al.*^{16,17}, Straub *et al.*¹⁸ and Netzel and co-workers¹⁹ have used picosecond spectroscopic studies of porphyrins and metalloporphyrins in an effort to gain a better understanding of the mechanisms of excited state relaxation. Irradiation of a metal-free porphyrin by 530-nm light (Fig. 2*a*) resulted in increased absorption at 590 nm by the excited state formed promptly after excitation. When excitation was performed with 355-nm light (Fig. 2*b*), bleaching of the Soret absorption band was observed, as a result of a population depletion of the ground state. Transient absorption spectroscopy shows that the excited state population persists for longer than 1 ns, which is consistent with the well-known strong fluorescence of metal-free porphyrins. The behaviour of the metalloporphyrin containing a closed-shell diamagnetic ion, Zn(II) protoporphyrin IX dimethyl ester, is analogous, as expected, to that of the metal-free porphyrin, namely the existence of an excited state lifetime of >1 ns.

Metalloporphyrins containing open-shell diamagnetic ions Ni(II), Pd(II) and Pt(II) were also investigated. The absorption behaviour of each of these three metalloporphyrins after excitation at 530 nm was found to vary in the following manner. For Pt(II) protoporphyrin IX dimethyl ester, 530-nm excitation

induces a long-lived (>1 ns) 'bleach', that is, ground-state depopulation, at 550 nm and a long-lived (>1 ns) absorption increase at 480 nm. As mentioned previously, Pt porphyrins are known to fluoresce only slightly or not at all; however, they phosphoresce strongly. Consequently, the long-lived component was attributed to the triplet state. The Pd(II) protoporphyrin IX dimethyl ester exhibited a biphasic decay of the excited state (Fig. 2*c*). The bleaching at 560 nm reveals a short-lived (~ 20 ps) component and a long-lived component which very slightly decreases during the 350-ps period of experimental observation. The corresponding kinetics were observed for excited-state absorbance increase at 480 nm. The short-lived component was assigned to the S_1 state and the long-lived component to the T_1 state. The kinetic behaviour of Ni(II) protoporphyrin IX dimethyl ester after 530-nm excitation is shown in Fig. 2*d, e*. The absorbance increase at 480 nm reveals a biphasic decay composed of a 10-ps fast component and ~ 250 -ps slow component. In contrast to the absorbance increase, the bleaching at 560 nm exhibits a single exponential recovery with a lifetime of 270 ps. This indicates the existence of a rapidly decaying excited state which does not repopulate the ground state (at 560 nm). Theoretical calculations^{20,21} predict the presence of low-lying S_d and T_d states of this metalloporphyrin. The state which does not decay to the ground state was assigned to the S_d state as neither phosphorescence nor fluorescence is observed for Ni(II) porphyrins in the visible range of

the spectrum and because the lifetime of the long-lived state is 250 ps.

Metalloporphyrins with an open-shell paramagnetic ion, such as Cu(II) or Ag(II), have also been studied. Biphasic kinetics for the bleaching process at 570 nm and also biphasic kinetics for the absorption of a new transient species at 480 nm were observed when Cu(II) protoporphyrin IX dimethyl ester was excited with 530-nm light. The time constant for the short-lived component of both the bleach and the transient absorption was found to be ~450 ps while that of the long-lived component of both was >1 ns. The excited states of this porphyrin have been described in terms of singdoublets (for example, 2S_1 and 2S_2), tripdoublets and quartets (for example, 2T_1 and 4T_1)¹³. As Cu(II) porphyrins do not fluoresce but are weakly phosphorescent, the following explanation for relaxation after photoexcitation was proposed on the basis of the picosecond absorption data. The 2S_1 state relaxes promptly (<6 ps) to the 2T_1 state via intersystem crossing. The first observable decay rate (lifetime $\approx$ 450 ps) results from an equilibrium established with the 4T_1 state which is slightly lower in energy than the 2T_1 state. The long-lived component, therefore, is responsible for the relaxation from the 4T_1 and 2T_1 states to the doublet ground state of the Cu porphyrin.

The metals within a metalloporphyrin clearly have a pronounced effect on the relaxation of the metalloporphyrin from excited states to the ground state and an effect on the radiationless transition process between states of different multiplicities, that is, intersystem crossing from a singlet to a triplet state is profoundly influenced by the heavy atom and the presence of unpaired electrons. Also, recent work by Reynolds *et al.*²² suggests that the pH significantly influences the optical spectrum of the transient species and the energy relaxation rates.

Dissociation of oxymyoglobin and carboxymyoglobin

The mechanisms of binding and dissociation of such ligands as oxygen and carbon monoxide by haemoglobin and myoglobin^{22,23} are still poorly understood, so our previously published^{24,25} picosecond spectroscopic investigations on the light-induced dissociation of myoglobin-CO and myoglobin-O₂ will not be the last of their kind.

Our first experiments showed that picosecond photo-induced dissociation of oxymyoglobin (MbO₂) and carboxymyoglobin (MbCO) occurs within picoseconds after excitation with a 530-nm picosecond pulse²⁴. While the exact mechanism of the dissociation and relaxation process is not known, the relaxation rate from the (π , π^*) excited state to the ground state is greatly enhanced by the presence of the metal atom as with other metalloporphyrins^{16,17,26}.

In our recent picosecond absorption study²⁵ we found that photodissociation of MbO₂ and MbCO was induced by excitation with a 355-nm or 532-nm picosecond pulse. Absorption changes were monitored at 440 nm (myoglobin absorption) and 420 nm (MbO₂ and MbCO bleaching) (Fig. 3). Excitation of MbO₂ or MbCO with 355- or 532-nm light induced a transient absorption at 440 nm within ~15–20 ps which persisted for at least 280 ps. Both MbO₂ and MbCO must have dissociated, as evidenced by the bleaching of their 420-nm bands, within 5–10 ps after excitation at either 355 or 532 nm. Subsequent decay of the 420- or 440-nm bands did not occur within the first 280 ps after excitation. An intermediate apparently is present which possesses a lifetime of ~6–8 ps. For picosecond excitation, the relative quantum yields of MbCO to that of MbO₂ were determined to be 3:1 for 355-nm excitation and 10:1 for 530-nm excitation (see Table 1 for a summary of these data).

With longer light pulses, the quantum yield for dissociation of MbCO was measured to be near unity²⁷ and that for MbO₂ to be ~0.03 (ref. 28). If direct dissociation alone accounted for energy dissipation, the dissociation of MbCO should have been 30 times faster than that of MbO₂, whereas it was essentially

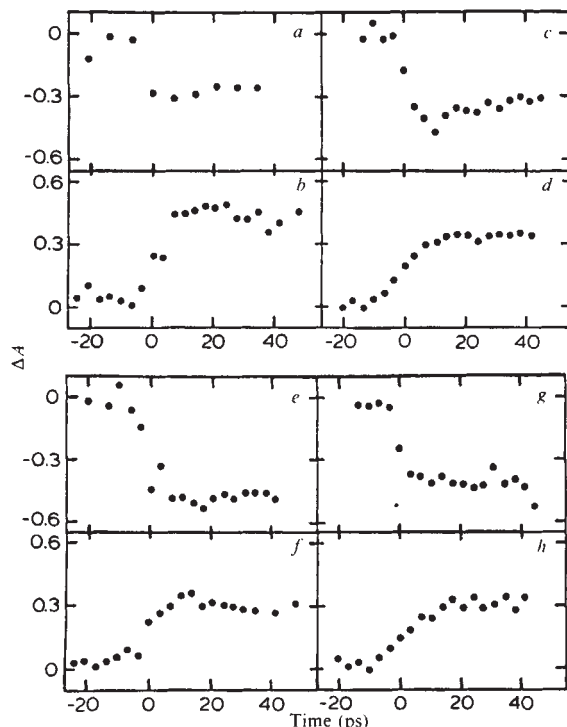


Fig. 3 Laser-induced absorption changes of MbO₂ with respect to time: a, 355-nm excitation, 420-nm interrogation; b, 355-nm excitation, 440-nm interrogation; c, 532-nm excitation, 420-nm interrogation; d, 532-nm excitation, 440-nm interrogation. Laser-induced absorption changes of MbCO with respect to time; e, 355-nm excitation, 420-nm interrogation; f, 355-nm excitation, 440-nm interrogation; g, 530-nm excitation, 420-nm interrogation; h, 530-nm excitation, 440-nm interrogation (from ref. 25).

equal for both when excited by 355-nm pulses of light²⁵. Clearly, another path for energy dissipation exists for MbO₂ which permits fast relaxation to either the ground or other non-dissociative state.

In the case of 532-nm excitation, the dissociation lifetimes of MbCO and MbO₂ were essentially the same and equal to those observed for 355-nm excitation. However, the ratio of the quantum yield for MbCO dissociation to that of MbO₂ dissociation had increased from 3 for 355-nm excitation to ~10 for 532-nm excitation.

$$\Phi_{\text{MbCO}}/\Phi_{\text{MbO}_2} = 3(\lambda_{\text{exc}} = 355 \text{ nm})$$

$$\Phi_{\text{MbCO}}/\Phi_{\text{MbO}_2} = 10(\lambda_{\text{exc}} = 532 \text{ nm})$$

To reconcile these data with the quantum yields obtained from longer pulse flash photolysis, again there must be more than one process involved in the dissociation of liganded myoglobin initiated by 532-nm excitation.

We have proposed²⁵ the following mechanism: MbL represents ground-state myoglobin with ligand, MbL* is its elec-

Table 1 Dissociation time constants and quantum efficiencies for MbO₂ and MbCO ligand detachment

	ϕ	τ dissociation (ps)			
		355-nm excitation		530-nm excitation	
		420-nm probe	440-nm probe	420-nm probe	440-nm probe
MbO ₂	0.03	3.5 ± 2	10.5 ± 2	3.5 ± 2	10 ± 1.5
MbCO	1	5 ± 2	12 ± 2.5	4.5 ± 2	10.5 ± 1.5
MbCO/MbO ₂	33	2.5 ± 0.5	3 ± 0.5	6.5 ± 2.5	9.5 ± 2
Δ_1 (MbCO-MbO ₂)		+1.5 ± 2	+1.5 ± 2	+1 ± 2	+0.5 ± 1.5

See ref. 25.

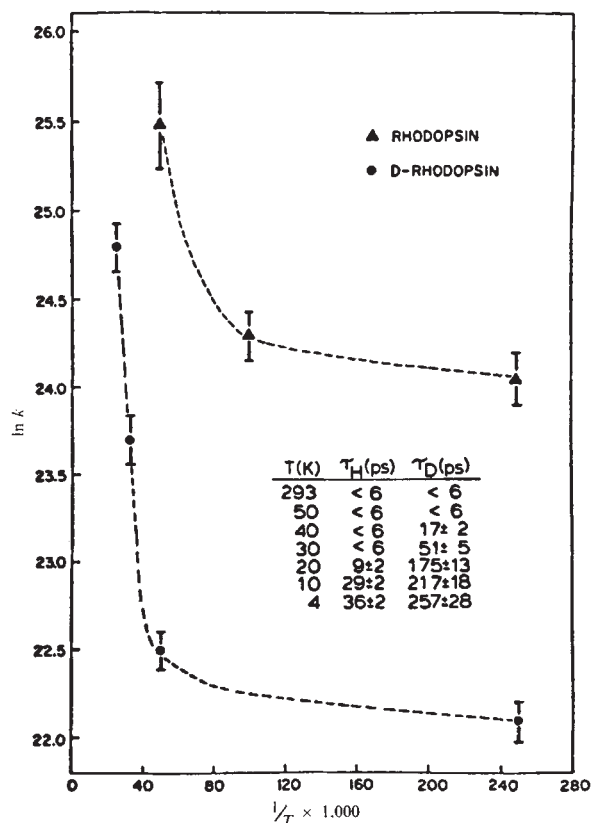
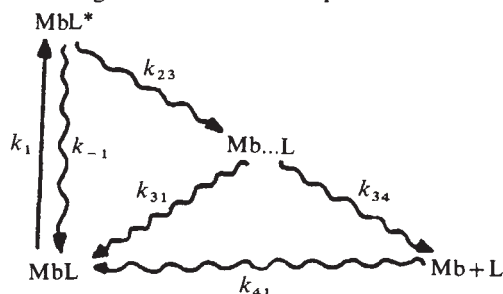


Fig. 4 Arrhenius plot, $\ln k$ (k in units of s^{-1}) for the formation of bathorhodopsin versus $1/T$ (T in units of K), of the listed kinetic data, where τ_H and τ_D are time constants for formation of native and deuterated bathorhodopsin, respectively (from ref. 34).

tronically excited state, $Mb \cdots L$ is the intermediate state, and $Mb + L$ are the ground states of the species after dissociation.



The rate constants k_{-1} , k_{31} , or k_{41} cannot account for the observed 10-ps lifetime because MbL disappears within ~ 5 ps and does not return for the next 280 ps. k_1 is the rate constant for absorption of a photon which, of course, is very rapid as evidenced by the rapid bleaching of the 420-nm band. Therefore, k_{23} , k_{34} or $(k_{23} + k_{34})$ are the processes responsible for the 10-ps kinetics.

While other, more involved, mechanisms are possible, we believe the simplest one to have $k_{34} \approx 10^{11} s^{-1}$ and $k_{31} \leq 10^{10} s^{-1}$. The quantum yield of dissociation is given by the ratios of k_{23}/k_t . This mechanism permits one to reconcile the observed quantum yields and kinetic data of oxymyoglobin and carboxymyoglobin dissociation.

Intermediates in visual transduction

Despite intense investigations of the process of vision, the primary sequence of events that are initiated by light absorbed by the visual chromophore, rhodopsin, is not well established. We will focus only on recent picosecond spectroscopic studies; for a recent extensive review of the subject the reader is referred to ref. 29.

Although absorption and Raman spectroscopy have provided a wealth of information on the initial process and the nature of the first intermediate, there is other controversy over the identification of the first intermediate—most commonly considered to be bathorhodopsin—and the process responsible for its formation. The primary changes in the absorption spectrum of rhodopsin were first observed at 77 K (refs 30–32) and were attributed to the first intermediate, bathorhodopsin. Using picosecond absorption spectroscopy, Busch *et al.*³³ first measured the room temperature kinetics of formation and decay of bathorhodopsin and showed that it was formed within 6 ps and decayed to the second intermediate, lumirhodopsin, with a time constant of ~ 30 ns.

The remarkably fast formation of bathorhodopsin prompted our low-temperature studies³⁴. The formation kinetics of bathorhodopsin were studied in optically clear glasses of the rhodopsin preparation in ethylene glycol over temperatures ranging from 300 to 4 K. Even at the lowest temperature studied, the risetime of the appearance of bathorhodopsin was found to be extremely fast, 36 ps.

Classically, the primary event in the formation of bathorhodopsin had been thought to be the *cis* $\rightarrow$ *trans* isomerization of the 11-*cis*-retinal moiety of rhodopsin to the all-*trans* form^{29,35,36}. Before our low temperature studies³⁴, the possibility of the occurrence of complete *cis* $\rightarrow$ *trans* isomerization of the retinal chromophore within 6 ps had been questioned^{33,37}. In view of the very fast formation of bathorhodopsin at 4 K, we proposed³⁴ proton translocation as an alternative mechanism at least at low temperature. The proposal was tested with samples of deuterated rhodopsin (D-rhodopsin) prepared by shaking the rhodopsin preparation with D_2O . In these conditions, only readily exchanged protons would be substituted with deuterium, and hydrogen atoms, for example, directly bonded to the hydrocarbon skeleton of the retinal moiety would be unaffected. A pronounced deuterium isotope effect was observed for the rate of formation of bathorhodopsin, $k_H/k_D = 7$. The temperature dependence of the rate of formation of bathorhodopsin also exhibited a non-Arrhenius behaviour and revealed that at very low temperatures the rate was practically independent of temperature. These data also revealed proton translocation, which occurs via quantum mechanical tunnelling at low temperatures. Analysis of the kinetic data permits the calculation of the energy barrier through which the proton tunnels and, therefore, the distance required for the proton translocation. The calculated distance of ~ 0.5 Å and the fact that the Schiff base proton was the only deuterium-exchangeable proton within the retinal moiety provided the basis for proposing that translocation of the Schiff base proton is at least a prominent component of the first event in vision³⁴.

Strong support for the hypothesis that *cis*–*trans* isomerization is the primary event in the visual transduction process seems to come from the belief that 11-*cis*-rhodopsin and 9-*cis*-rhodopsin lead to a common bathorhodopsin. For example, Ottolenghi²⁹ concluded that the first event in the visual process is *cis*–*trans* isomerization at the 11-*cis* double bond of the retinal skeleton on the ground, that photoexcitation of both 11-*cis*-rhodopsin and 9-*cis*-rhodopsin lead to the formation of a common intermediate which must possess an all-*trans* geometry. Raman experiments³⁸, data presented by Rosenfeld *et al.*³⁶, Green *et al.*³⁹ and Monger *et al.*⁴⁰ and several theoretical models^{37,41} support the proposal of a common all-*trans* intermediate and seem to provide a strong argument for *cis*–*trans* isomerization and not merely a slight skeletal deformation, as was proposed by Busch *et al.*³³ and Peters *et al.*³⁴, as the process generating bathorhodopsin.

However, our recent picosecond absorption study⁴² suggest that bathorhodopsin formed from photoexcitation of 9-*cis*-rhodopsin is not the same as the bathorhodopsin generated from 11-*cis*-rhodopsin. Using a 25-ps, 532-nm pulse generated by a KDP crystal from the fundamental 1,064-nm pulse of a mode-locked Nd^{3+} :YAG laser for excitation of 9-*cis*- and 11-*cis*-rhodopsin, we obtained difference absorption spectra

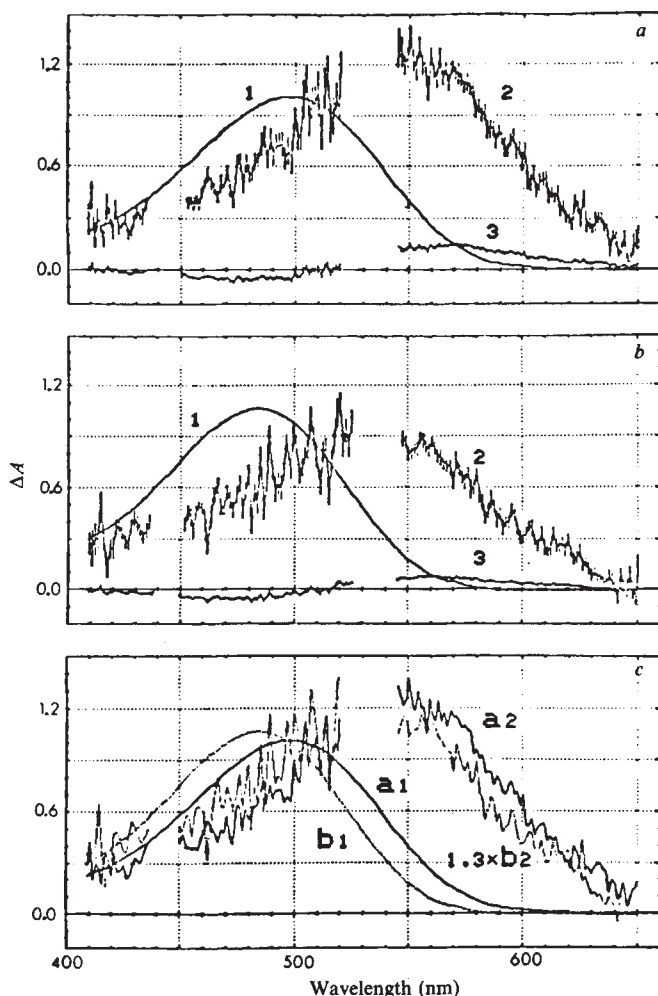


Fig. 5 *a*, (1) 11-*cis*-rhodopsin spectrum; (2) calculated spectrum of bathorhodopsin from 11-*cis*-rhodopsin (resolved from difference spectrum (curve 3) by assuming a 15% bleach); (3) difference spectrum. *b*, (1) 9-*cis*-rhodopsin spectrum; (2) calculated spectrum of bathorhodopsin from 9-*cis*-rhodopsin (resolved from difference spectrum (curve 3) by assuming a 13% bleach); (3) difference spectrum. *c*, Comparison of the two calculated bathorhodopsin spectra (*a*₂ and *b*₂). Both are scaled to the same maximum absorption to illustrate the spectral shift. *a*₁ and *b*₁ are the respective 11-*cis*- and 9-*cis*-rhodopsin spectra (from ref. 42).

with an interrogating picosecond continuum. The probing time which was 85 ps after excitation of the rhodopsin sample was sufficient to permit excited-state rhodopsin to decay to bathorhodopsin (time constant of ~6 ps) and orders of magnitude shorter than the decay of bathorhodopsin to the next intermediate, lumirhodopsin (time constant of ~30 ns). Therefore, the absorption spectra measured correspond to bathorhodopsin. The difference spectra were calculated according to

$$\Delta A(\lambda) = -\log_{10} \left(\frac{\left(\frac{S^{\text{ex}}(\lambda)}{R^{\text{ex}}(\lambda)} \right)}{\left(\frac{S^{\text{noex}}(\lambda)}{R^{\text{noex}}(\lambda)} \right)} \right)$$

where ΔA is the change in absorbance, $S^{\text{ex}}(\lambda)$ and $R^{\text{ex}}(\lambda)$ are the intensities of the sample probe and of the reference probe, respectively, as a function of wavelength, when the sample is subjected to photoexcitation by the laser pulse. $S^{\text{noex}}(\lambda)$ and $R^{\text{noex}}(\lambda)$ are the intensities of the sample probe and of the reference probe, respectively, as a function of wavelength when the sample is not subjected to photoexcitation by the laser pulse. Low excitation energies were used to insure that difference spectra were not artefacts resulting from saturation of the electronic transition or biphotonic events.

Figure 5 illustrates the difference spectra obtained 85 ps after excitation of 9-*cis*-rhodopsin and 11-*cis*-rhodopsin at 290 K.

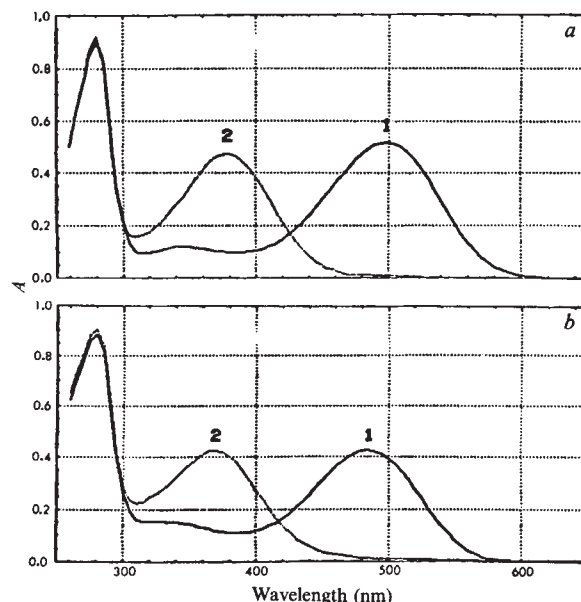


Fig. 6 Absorption spectra of 11-*cis*-rhodopsin (*a*) and 9-*cis*-rhodopsin (*b*) in ammonyx-LO. Curves labelled (1) were recorded before irradiation with white light while those labelled (2) were recorded after irradiation with white light for 10 min (from ref. 42).

Two important observations not previously noted are quite evident on inspection of the two spectra. First, the absorption and bleached bands of the difference spectrum resulting from 9-*cis* excitation and its isosbestic point are shifted to longer wavelengths by ~10 nm relative to the corresponding positions in the 11-*cis* difference spectrum. This shift is similar to that observed for the ground-state absorptions of 11-*cis*- and 9-*cis*-rhodopsin (Fig. 6). Second, while both 11-*cis*-rhodopsin and 9-*cis*-rhodopsin exhibit the same amount of maximum bleach (absorption decrease), the ratio of their absorption maxima is not 1:1 but 1.5:1.

The absorption spectra (Fig. 5) of bathorhodopsins formed 85 ps after excitation of 11-*cis*-rhodopsin and 9-*cis*-rhodopsin can be obtained by subtracting the bleach from the difference spectrum. Scaling of the absorption bands clearly illustrates that they are not superimposable. This spectral difference persists up to 8 ns after excitation of the sample. These data clearly show that a common intermediate is not formed when 9-*cis*-rhodopsin and 11-*cis*-rhodopsin are excited with 532-nm light.

Although these data do not record or examine skeletal deformations, that is, *cis-trans* isomerization, they do cast considerable doubt on the premise of an 11-*cis* and 9-*cis* common batho-intermediate. Further evidence against that proposal comes from recent Raman data⁴³ which show that there is little configurational change. IR spectra⁴⁴ of bacteriorhodopsin, which contains essentially the same chromophore as rhodopsin, proves that translocation of the proton attached to the Schiff base can take place. Although our picosecond absorption data^{33,34} could not distinguish and positively identify the proton involved in the translocation, we proposed that the proton attached to the Schiff base would translocate and further that it would move towards the Schiff base nitrogen. Studies by Sandorfy and co-workers^{45,46} suggest that this proposal is correct. The recent IR⁴⁴ studies indicate that the Schiff base proton of bacteriorhodopsin translocates away from the counter ion, supporting our original proposal.

Concluding remarks

The examples presented here are supposed to illustrate the clarity with which intricate step-by-step processes can be resolved with the aid of mode-locked picosecond laser pulses. It is rather gratifying that picosecond spectroscopy can shed new

knowledge and suggest new mechanisms not only for relatively simple molecular processes but also in complicated biological systems. Although sometimes controversial when first proposed, these mechanisms have withstood the test of time and, after confirmation by other methods, have become accepted. Currently, a strong effort is being made to improve picosecond Raman spectroscopy and to generate high-quality short pulses tunable from the vacuum UV into the IR. With ever-improving

technological advances, presently impenetrable mechanisms will be clarified by means of absorption, emission and Raman spectroscopy which utilize extremely short light pulses. These techniques will allow the complete study of a chemical or biological reaction and facilitate the advance of picosecond spectroscopy from the two-wavelength relaxation method of 1968⁴⁷ to a most versatile, time-resolved, spectroscopic technique.

1. Rentzepis, P. M. *Science* **202**, 174–182 (1978).
2. Reynolds, A. H. & Rentzepis, P. M. *Adv. bot. Res.* **8**, 1–23 (1980).
3. Ippen, E. P. & Shank, C. V. in *Ultrashort Light Pulses* (ed. Shapiro, S. L.) (Springer, Heidelberg, 1977).
4. Shapiro, S. L., Cavanaugh, R. R. & Stephenson, J. C. *Opt. Lett.* **6**, 470–474 (1981).
5. Eastwood, D. & Gouterman, M. *J. molec. Spectrosc.* **30**, 437–458 (1969).
6. Tsvirko, M. P., Stelmakh, G. F., Pyatosin, V. E., Solovyov, K. N. & Kachura, T. F. *Chem. Phys. Lett.* **73**, 80–83 (1980).
7. Magde, D., Windsor, M. W., Holten, D. & Gouterman, M. *Chem. Phys. Lett.* **29**, 182–187 (1974).
8. Gouterman, M., Mathies, R. A., Smith, B. E. & Caughey, W. S. *J. chem. Phys.* **52**, 3795–3802 (1979).
9. Smith, B. E. & Gouterman, M. *Chem. Phys. Lett.* **2**, 517–519 (1968).
10. Becker, R. S. & Kasha, M. *J. Am. chem. Soc.* **77**, 3669–3770 (1955).
11. Treibs, A. *Ann. N.Y. Acad. Sci.* **206**, 97–115 (1973).
12. Gouterman, M. *Ann. N.Y. Acad. Sci.* **206**, 70–83 (1973).
13. Ake, R. L. & Gouterman, M. *Theor. chim. Acta* **15**, 20–42 (1969).
14. Gouterman, M., Schwarz, F. P., Smith, P. D. & Dolphin, D. *J. chem. Phys.* **59**, 676–690 (1973).
15. Corwin, A. H., Chivvis, A. B., Poor, R. W., Whitten, D. G. & Baker, E. W. *J. Am. chem. Soc.* **90**, 6577–6583 (1968).
16. Kobayashi, T., Huppert, D., Straub, K. D. & Rentzepis, P. M. *J. chem. Phys.* **70**, 1720–1726 (1979).
17. Kobayashi, T., Straub, K. D. & Rentzepis, P. M. *Photochem. Photobiol.* **29**, 925–931 (1979).
18. Straub, K. D., Huppert, D. & Rentzepis, P. M. *J. Photochem.* **17**, 128–129 (1981).
19. Bergkamp, M. A., Dalton, J. & Netzel, T. L. *J. Am. chem. Soc.* **104**, 253–259 (1982).
20. Ake, R. L. & Gouterman, M. *Theor. chim. Acta* **17**, 408–416 (1970).
21. Adamczyk, A. & Wilkinson, F. *JCS Faraday II* **68**, 2031–2041 (1972).
22. Reynolds, A. H., Milton, S. V., Straub, K. D. & Rentzepis, P. M. *Biophys. J.* (Submitted).
23. Antonini, E. & Brunori, M. *Front. Biol.* **21**, 1–445 (1971).
24. Noe, L. J., Eisert, W. G. & Rentzepis, P. M. *Proc. natn. Acad. Sci. U.S.A.* **75**, 573–577 (1978).
25. Reynolds, A. H., Rand, S. D. & Rentzepis, P. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2292–2296 (1981).
26. Huppert, D., Straub, K. D. & Rentzepis, P. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4139–4143 (1977).
27. Bücher, T. & Kaspers, J. *Biochim. biophys. Acta* **1**, 21–34 (1947).
28. Gibson, O. H. & Ainsworth, S. *Nature* **180**, 1416–1417 (1957).
29. Ottolenghi, M. *Adv. Photochem.* **12**, 97–200 (1980).
30. Yoshizawa, T. & Kito, Y. *Nature* **182**, 1604–1605 (1958).
31. Grellman, K. H., Livingston, R. & Pratt, D. *Nature* **193**, 1258–1260 (1962).
32. Yoshizawa, T. & Wald, G. *Nature* **197**, 1279–1286 (1963).
33. Busch, G. E., Applebury, M. L., Lamola, A. A. & Rentzepis, P. M. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2802–2806 (1972).
34. Peters, K., Applebury, M. L. & Rentzepis, P. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3119–3123 (1977).
35. Wald, G. *Nature* **219**, 800–807 (1968).
36. Rosenfeld, T., Honig, B., Ottolenghi, M., Hurley, J. & Ebrey, T. G. *Pure appl. Chem.* **49**, 341–351 (1977).
37. Warshel, A. *Nature* **260**, 679–683 (1976).
38. Mathies, R., Oseroff, A. R. & Stryer, L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1–5 (1976).
39. Green, B. H., Monger, T. G., Alfano, R. R., Aton, B. & Callender, R. H. *Nature* **269**, 179–180 (1977).
40. Monger, T. G., Alfano, R. R. & Callender, R. H. *Biophys. J.* **27**, 105–115 (1979).
41. Birge, R. R. & Hubbard, L. M. *J. Am. chem. Soc.* **102**, 2195–2205 (1980).
42. Spalink, J. D., Reynolds, A. H., Rentzepis, P. M., Applebury, M. L. & Sperling, W. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
43. Braiman, M. & Mathies, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 403–407 (1982).
44. Rothschild, K. J. & Marrero, H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4045–4049 (1982).
45. Favrot, J., Leclercq, J. M., Roberge, R., Sandorfy, C. & Vocelle, D. *Chem. Phys. Lett.* **53**, 433–438 (1978).
46. Favrot, J., Sandorfy, C. & Vocelle, D. *Photochem. Photobiol.* **28**, 271–272 (1978).
47. Rentzepis, P. M. *Chem. Phys. Lett.* **2**, 117–120 (1968).

ARTICLES

Low-frequency variability and predicted superluminal motion in 3C147

R. S. Simon^{*†}, A. C. S. Readhead^{*}, A. T. Moffet^{*}, P. N. Wilkinson[‡], B. Allen[§] & B. F. Burke[§]

^{*} Owens Valley Radio Observatory, California Institute of Technology, Pasadena, California 91125, USA

[†] E.O. Hulburt Center for Space Research, Naval Research Laboratory, Washington DC 20375, USA

[‡] Nuffield Radio Astronomy Observatory, Jodrell Bank, Manchester SK11 9DL, UK

[§] Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

VLBI observations of 3C147 reveal that its core is a low-frequency variable radio source which has brightened by a factor of 2 in 6 years. In combination with X-ray observations, this implies that bulk relativistic motion is taking place within the core, and leads to the prediction that 3C147 is a member of the class of superluminal radio sources.

LOW-FREQUENCY variability in extragalactic radio sources poses some theoretical problems^{1,2} as the time scales involved imply brightness temperatures which exceed the inverse Compton limit³ of 10^{12} K. One possible solution is that the radiation regions are moving towards us with a relativistic bulk speed, but this may lead to very large values of total energy, or require special source geometries^{2,4}. We report here the first direct observations of component variability in an extragalactic radio source at low frequency. An analysis of the radio and X-ray observations reveals that at least part of the radio emission region must be moving towards us relativistically with a Doppler factor, $\delta \geq 6$.

We have observed the quasar 3C147 (0538+49, $z = 0.545$) using very long baseline interferometry (VLBI) at 329 MHz at two different epochs. The first observations were made in March 1975 using a three-station interferometer consisting of Jodrell

Bank, Cheshire, UK (JBNK); Green Bank, West Virginia, USA (NRAO); and Big Pine, California, USA (OVRO). These observations have been discussed elsewhere⁵. In February 1981, we re-observed 3C147 with a seven-station interferometer including the above three stations and four additional stations in the US: Haystack, Massachusetts; North Liberty, Iowa; Fort Davis, Texas; and Hat Creek, California. A full discussion of the details of this second observation is in preparation⁶. Hybrid maps⁷ made from the two data sets are shown in Fig. 1. The 1981 map has been convolved with the beam appropriate to the 1975 map, and both maps have the same absolute contour levels so that they may be compared directly. Each map can be described roughly as showing an unresolved core with a partially resolved jet consisting of two knots extending to the south-west. Due to the very low dynamic range of the 1975 map, only the positions and the relative brightnesses of the

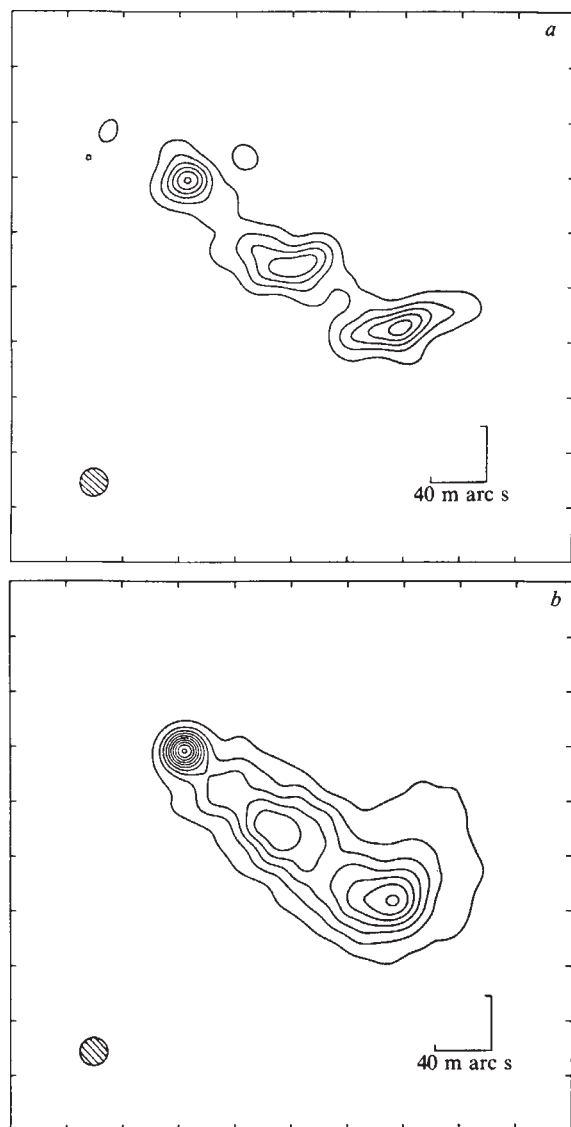


Fig. 1 Hybrid maps of 3C147 at 329 MHz. In each map, the lowest contour corresponds to a brightness temperature of 3×10^9 K, and the contour interval is 6×10^9 K. The δ functions which compose each map have been convolved with a circular gaussian restoring beam (lower left) of 20 m arc s diameter. *a*, Map from March 1975. The core (the most northeasterly of the three components) is only slightly greater in peak brightness temperature than the 2 knots in the jet. *b*, Map from February 1981. The core is substantially brighter than it was in 1975.

three components are well determined⁵. The shapes of the components in the 1975 map are poorly defined and in particular, the width of the jet is not strongly constrained. In contrast, the 1981 map is very reliable with a dynamic range exceeding 25:1.

Comparisons between the two maps reveal that the barely resolved core has increased in flux density by a factor of ~ 2 in the 6 yr between observations, from 0.9 Jy in 1975 to 2.0 Jy in 1981. This is the first direct observation of a change in the flux density of a component of an extragalactic radio source below a frequency of 2 GHz, as well as the first positive detection of low-frequency variability in 3C147. While Fisher and Erickson⁸ found 3C147 to "show evidence" of variability below 1 GHz, Fanti *et al.*⁹ failed to detect 3C147 as a low-frequency variable, with an upper limit to any variations at 408 MHz of $\sim 5\%$ of the total flux density.

Readhead *et al.*¹⁰ first pointed out that there is an advantage in using VLBI to study low-frequency variability, as the visibility on the largest baseline would be sensitive only to the most

compact, and presumably varying, structure. In the present case the advantage of VLBI over single dish observations is clear, since the observed flux density increase of 1 Jy amounts to only 2% of the total source intensity, but corresponds to a factor of 3 increase in the correlated flux density on the longest baseline (JBNK–OVRO).

Despite their use of VLBI, Readhead *et al.* detected no appreciable variation in structure in 3C147 between December 1973 and March 1975 at 609 MHz, and succeeding observations at this frequency have been marred by repeated failures at one or more telescopes. Their failure to detect variation interferometrically is possibly due to three factors. First, the long time constant of the variation requires observations spaced by several years. Second, a reduction in the amplitude of variability at higher frequencies is implied, with the spectral index of the variable component ~ 1 (the flux density S_ν at frequency ν is related to the spectral index α by $S_\nu \propto \nu^{-\alpha}$). This is consistent with an observed property of other low-frequency variables, namely, the amplitude of variability decreases as the observing frequency increases¹¹. Finally, variability in the core may occur intermittently.

The evidence for variability in the source can be seen directly in the closure phase data. Figure 2 shows part of the closure phase measured for the triangle JBNK–NRAO–OVRO for both 1975 and 1981. As the closure phase is free of both calibration and systematic errors (to approximately the 2° level), the large differences between the two data sets (up to $\sim 50^\circ$) imply significant changes in the source structure. At 16.2 h GST the closure phase for both observations undergoes a large phase jump, corresponding to a deep minimum in the visibility amplitude on the baseline NRAO–OVRO. The 1981 data undergo the larger phase change because the amplitude minimum was deeper in 1981 than in 1975. This was caused by the stronger beating (in 1981 compared with 1975) of the core against the knot at the end of the jet, due to the increased flux density of the core. Detailed modelling⁶ indicates that the core has increased in brightness by a factor of 2.2 ± 0.2 .

In addition to the low-frequency variability, the 10.6-GHz flux density has been slowly decreasing over the past decade¹², which is a further indication of activity within the source.

Relativistic bulk motion

The direct evidence for relativistic bulk motion comes from a comparison of the predicted inverse Compton X-ray emission with the observed emission, and the assumption that the observed X-ray flux density is an upper limit to the inverse Compton X-ray flux density. Marscher and Broderick were the first to apply this method of analysis in detail^{13,14}, and successfully predicted superluminal motion in the quasar NRAO 140¹⁵. To perform the analysis and make this comparison it is first necessary to determine the spectrum shape and the size of the radio emitting region.

Readhead and Wilkinson¹⁶ made a hybrid map of 3C147 at 1,671 MHz in September 1976. We have combined this map with maps at 329 MHz (ref. 5) and 609 MHz (ref. 17) from March 1975, the 2.3-GHz map of Phillips and Shaffer¹⁸, and included more limited data at 2.7 GHz, 15 GHz and 86 GHz (see ref. 5 and references therein). This allows us to determine the spectrum of the core (see Fig. 3). Note that the 86-GHz core flux density is based on the total flux density minus the flux density extrapolated for the jet. The observations of the core show that it is optically thick at 329 MHz and that the spectrum has a maximum between 609 and 1,671 MHz. We have fitted the spectrum to that of a spherical, homogeneous, self-absorbed synchrotron source. Such a spectrum is completely specified by three parameters: α , the high-frequency (optically thin) spectral index; and ν_m and S_m , the frequency and flux density, respectively, at the maximum in the spectrum. Note that we define S_m as the flux density at ν_m extrapolated from the optically thin spectrum. For the core of 3C147, we find that $\alpha = 0.45$, $\nu_m = 0.8$ GHz, and $S_m = 4.7$ Jy (the actual peak flux density is 4.0 Jy). The estimated errors for these

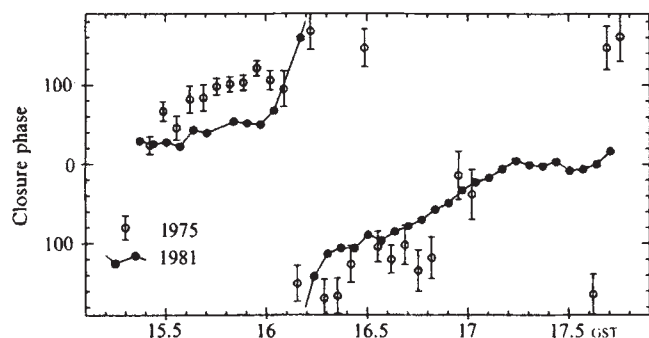


Fig. 2 Closure phase observations on the triangle JBNK-NRAO-OVRO at 329 MHz from 1975 and 1981. The 1975 data are plotted as open circles with error bars; the 1981 data as filled circles (the 1981 data are connected with a solid line for clarity). The error bars for the 1981 data are approximately the same size as the symbol used.

parameters are $\sim 10\%$, and do not significantly affect the calculations which follow. Furthermore, these three parameters are not independent; in the calculations below, an error in the determination of one of these parameters will be partially compensated by changes in the values of the other parameters (see Fig. 3, and equation (1); equation (1) is insensitive to α , while S_m and ν_m are roughly proportional and vary monotonically with α).

We have followed the standard procedures¹⁹ for calculating the inverse Compton X-ray emission for the radio emitting region. For an assumed uniform, spherical, synchrotron source the Doppler factor, $\delta = [\gamma(1 - \beta \cos \phi)]^{-1}$ is given by^{20,21}

$$\delta \geq 0.27\theta^{-1.59} S_x^{-0.20} S_m \nu_m^{-1.30} \quad (1)$$

for $\alpha = 0.45$, where θ is the angular diameter (m arcs), S_x is the 1.3-keV X-ray flux density (Jy), and ν_m is in GHz. This holds for a redshift $z = 0.545$, and assumes that the optically thin spectrum is straight to a frequency $\nu_2 \sim 100$ GHz. The observed Einstein X-ray flux²² of 3C147 in the 0.27–3.28-keV band is $2.9 \times 10^{-13} \text{ erg s}^{-1} \text{ cm}^{-2}$. After correction for galactic absorption this corresponds to an X-ray flux density of $6.5 \times 10^{-8} \text{ Jy}$ at 1.29 keV. Substituting this value and the appropriate values derived from the radio spectrum we find

$$\delta \geq 50\theta^{-1.59} \quad (2)$$

This relationship depends on the assumptions of source geometry and negligible absorption of the inverse Compton X rays, and is relatively insensitive to the high-frequency spectral index in the range $0.25 < \alpha < 0.75$. The overall angular size of the core is 5 m arcs, thus $\delta \geq 3.6$. Note that this value is independent of the distance to 3C147. The core consists of at least two components^{16,23}, with angular sizes of ≥ 2.5 and ~ 3.5 m arcs. Thus this is a strong lower limit to the Doppler factor. Because the individual spectra of the components within the core are unknown, we cannot improve this particular limit to δ .

An additional constraint on δ and θ can be derived from the variability time scale t_v , if causality is not to be violated²⁰.

$$\theta \leq 0.13 t_v \delta r^{-1} \text{ m arcs} \quad (3)$$

where r is the co-moving coordinate distance in Gpc, and (following Burbidge *et al.*²) the variability time scale t_v at 329 MHz is defined by

$$t_v = |d(\ln S_{329})/dt|^{-1} \quad (4)$$

In the present case we have $t_v = 7.5$ yr. The co-moving coordinate distance is 2.0 Gpc assuming $H_0 = 60 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0.5$. From equations (3) and (4) we derive a second expression for δ which, again, depends only on θ :

$$\delta \geq 2.1\theta \quad (5)$$

Taken together, and allowing for errors in the measurement of parameters for the core, equations (2) and (5) imply that $\delta \geq 7$

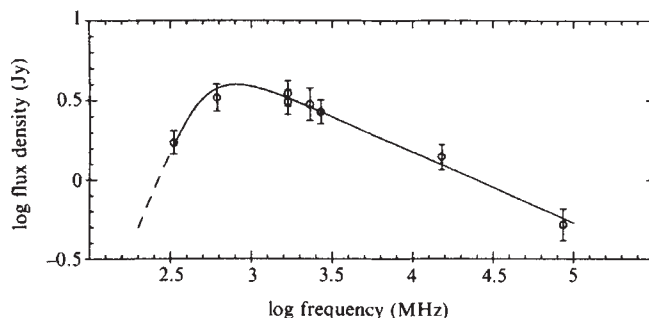


Fig. 3 The radio spectrum of the core of 3C147. Measurements are represented as open circles with error bars. The solid line is the best fit of a model (in a least-squares sense) to the data. The model consists of a homogeneous, spherical, incoherent synchrotron source.

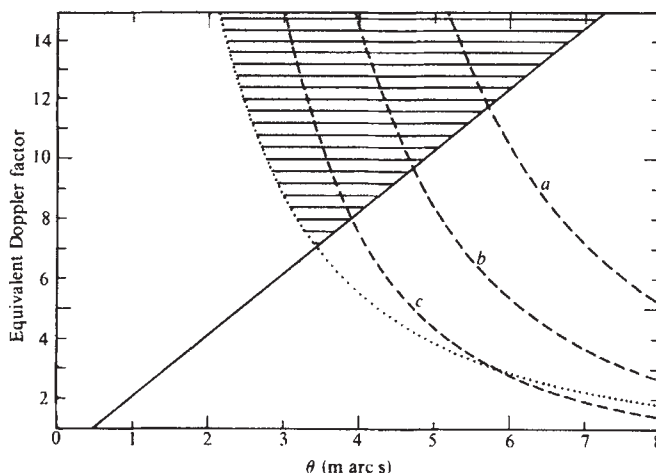


Fig. 4 Limits on the equivalent Doppler factor δ and the angular diameter θ for the core of 3C147. The dotted line is a lower limit to δ based on the requirement that the observed X-ray flux density cannot exceed the predicted inverse Compton X-ray flux density. The solid line is an upper limit to θ based on the variability time scale observed for the core, and assuming the cosmological distance to the source. The hatched area represents the region of allowed values of δ and θ . The three dashed lines are the curves of θ plotted against δ for three different values of u_e/u_m : a, $u_e/u_m = 0.01$ (field-dominated); b, $u_e/u_m = 1$ (equipartition); c, $u_e/u_m = 100$ (particle-dominated).

(see Fig. 4). It is therefore clear that there is bulk relativistic motion towards us in the core of 3C147.

Superluminal motion is likely as for $\delta > 7$, the angle to the line of sight must be $\phi \leq 8^\circ$, and for a random distribution of ϕ , the most probable observed velocity is²⁴ $v/c(\text{apparent}) \sim \delta(1+z) \sim 9$. VLBI observations were made in October 1982 to look for superluminal motion in 3C147. These have not yet been analysed.

An expansion or separation of two components in the core at this apparent speed would be readily observable. The corresponding increase in size is ~ 2.0 m arcs in 6 yr (for $v/c = 10$, $\mu \approx 0.3 \text{ m arcs yr}^{-1}$). This is many times the estimated error in measuring either the separation between the core components or the overall angular size of the core by VLBI. Failure to detect significant motion in the core would imply either that $\phi \ll 10^\circ$, so that $v/c \ll \delta(1+z)$, or that the dominant emission mechanism in the core is not incoherent synchrotron radiation. In the first case, the angle to the line of sight would be $\phi \leq 3^\circ$. In either case, if proper motion in the core in 6 yr is less than the minimum detectable (≈ 0.2 m arcs at 1.6 GHz), then $v/c(\text{apparent}) \leq 1$. Therefore the October 1982 observations will be a good test of our prediction of superluminal motion in the core of 3C147.

A weak test of this prediction can be made, based on existing VLBI observations at two frequencies. In the September 1976

18-cm map of Readhead and Wilkinson, the core had a partially resolved double structure with a separation between components of 2.6 m arcs. In January 1981 it was observed with VLBI at 6 cm (ref. 23); the core had a separation of ~ 3.6 m arcs, corresponding to a possible proper motion of $\mu \approx 0.23$ m arcs yr^{-1} . Including the time dilation correction factor due to the redshift, this implies an apparent separation velocity of $v/c \sim 7$. This is suggestive, but not conclusive, as the possibility exists that the different separations could be due to strong spectral gradients across the individual components in the core, rather than motion of the components.

Energy

In spite of the bulk relativistic motion occurring in the core of 3C147, an extreme value of the total energy is not required. From standard synchrotron formulae^{2,19} we can estimate the energy densities in the relativistic electrons (u_e) and in the magnetic field (u_M), in the rest frame of the emitting region. Due to the complex structure of the core, we shall assume $\theta = 5$ m arcs and $\delta = 10$, these being the most conservative values for these parameters consistent with the observations and with the constraints on δ and θ . We find that $u_e \sim 4 \times 10^{-8}$ erg cm^{-3} and $u_M \sim 1.4 \times 10^{-7}$ ergs cm^{-3} . The total energy of the core is then $U_T \sim 10^{53}$ erg and $u_e/u_M \sim 0.3$. In addition, for the core of 3C147 the following relationship for the ratio u_e/u_M applies:

$$\frac{u_e}{u_M} \sim 1.3 \times 10^{10} \theta^{-17} \nu_m^{-17} S_m^8 \delta^{-7} (1+z)^8 r^{-1} \quad (6)$$

Figure 4 shows the relationship between θ and δ for three different values of u_e/u_M . We see that $u_e/u_M \sim 1$ (within ~ 2 orders of magnitude) is allowed by the constraints on θ and δ .

Therefore, within the rest frame of the radiating region the core could be near equipartition of energy density between the magnetic field and the relativistic electrons (see ref. 25). In contrast, lower frequency observations²⁵ have shown that the large scale structure in 3C147 (~ 0.5 arcs, or ~ 3 kpc) is near

equipartition, but only if no bulk relativistic motion is assumed for that structure.

The total energy we have found can be compared with the total energy required for the radio core of 3C345, for which $U_T \sim 10^{56}$ erg (ref. 24), or the even more extreme possibilities discussed by Burbidge *et al.*²

Implicit in the derivation of the present relatively reasonable value of the total energy is the assumption that any relativistic motion which is occurring is in the form of component motion nearly along the line of sight, and not in the form of spherical expansion. Furthermore, we have assumed that any contribution to the total energy from non-relativistic matter swept up or entrained by the moving components of the core is negligible or, equivalently, that the motion is occurring in a channel of negligible density. This model is essentially the relativistic beaming model discussed by Readhead *et al.*²⁶ and others²⁷⁻³⁰ to explain asymmetric core radio sources.

A self-consistent picture of 3C147 as a low-frequency variable has emerged which is consistent with the available observations, yet does not require an extreme value of the total energy. The key to understanding both the rapid flux density increase of the core at 329 MHz and the strength of the observed X-ray flux density is the occurrence of bulk relativistic motion within the core. Together with NRAO 140 there are now two extragalactic low-frequency variable radio sources in which bulk relativistic motion has been demonstrated. Thus it may well be that bulk relativistic motion is responsible for the required time scales in most, if not all, low-frequency variables.

We thank the US VLBI Network and the observers and technical staff of all the network observatories involved in this work, in particular, R. L. Mutel at North Liberty Radio Observatory who built an entire receiving system for these observations, also R. D. Blandford and S. C. Unwin for helpful comments. This work was supported by the NSF via grant AST 79-13249 to the Owens Valley Radio Observatory. The National Radio Astronomy Observatory is operated by Associated Universities Inc., under contract to the NSF.

Received 23 September 1982; accepted 7 January 1983.

1. Jones, T. W. & Burbidge, G. R. *Astrophys. J.* **186**, 791-799 (1973).
2. Burbidge, G. R., Jones, T. W. & O'Dell, S. L. *Astrophys. J.* **193**, 43-54 (1974).
3. Kellermann, K. I. & Pauliny-Toth, I. I. K. *Astrophys. J. Lett.* **155**, L71-L78 (1969).
4. Blandford, R. D. & McKee, C. F. *Mon. Not. R. astr. Soc.* **180**, 343-371 (1977).
5. Simon, R. S., Readhead, A. C. S., Moffet, A. T., Wilkinson, P. N. & Anderson, B. *Astrophys. J.* **236**, 707-713 (1980).
6. Simon, R. S. *et al.* Preprint (California Institute of Technology).
7. Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J. Lett.* **223**, 25-36 (1978).
8. Fisher, J. R. & Erickson, W. C. *Astrophys. J.* **242**, 884-893 (1980).
9. Fanti, C. *et al.* *Astr. Astrophys. Suppl.* **45**, 61-78 (1981).
10. Readhead, A. C. S., Wilkinson, P. N. & Purcell, G. H. *Astrophys. J. Lett.* **215**, L13-L15 (1977).
11. Spangler, S. R. & Cotton, W. D. *Astr. J.* **86**, 730-746 (1981).
12. Andrew, B. H., MacLeod, J. M. & Feldman, P. A. *Astr. Astrophys.* **99**, 36-38 (1981).
13. Marscher, A. P. & Broderick, J. J. *Astrophys. J. Lett.* **247**, L49-L52 (1981).
14. Marscher, A. P. & Broderick, J. J. *Astrophys. J.* **249**, 406-414 (1981).
15. Marscher, A. P. & Broderick, J. J. *IAU Symp.* No. 97, 359-360 (1982).

16. Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J.* **235**, 11-17 (1980).
17. Wilkinson, P. N., Readhead, A. C. S., Purcell, G. H. & Anderson, B. *Nature* **269**, 764-768 (1977).
18. Phillips, R. B. & Shaffer, D. B. *Astrophys. J.* (submitted).
19. Jones, T. W., O'Dell, S. L. & Stein, W. A. *Astrophys. J.* **192**, 261-278 (1974).
20. Marscher, A. P. *et al.* *Astrophys. J.* **233**, 498-503 (1979).
21. Marscher, A. P. *Astrophys. J.* **264**, 296-297 (1983).
22. Zamorani, G. *et al.* *Astrophys. J.* **245**, 357-374 (1981).
23. Preuss, E., Alef, W., Pauliny-Toth, I. & Kellermann, K. I. *IAU Symp.* **97**, 289-290 (1982).
24. Unwin, S. C. *et al.* *Astrophys. J.* (in the press).
25. Scott, M. A. & Readhead, A. C. S. *Mon. Not. R. astr. Soc.* **180**, 539-550 (1977).
26. Readhead, A. C. S., Walker, R. C., Pearson, T. J. & Cohen, M. H. *Nature* **276**, 768-771 (1978).
27. Scheuer, P. A. G. & Readhead, A. C. S. *Nature* **277**, 182-185 (1979).
28. Blandford, R. D. & Konigl, A. *Astrophys. J.* **232**, 34-48 (1979).
29. Marscher, A. P. *Astrophys. J.* **239**, 296-304 (1980).
30. Readhead, A. C. S., Hough, D. H., Ewing, M. S., Walker, R. C. & Romney, J. D. *Astrophys. J.* **265**, 107-131 (1983).

Infectious vaccinia virus recombinants that express hepatitis B virus surface antigen

Geoffrey L. Smith, Michael Mackett & Bernard Moss

National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA

Potential live vaccines against hepatitis B virus have been produced. The coding sequence for hepatitis B virus surface antigen (HBsAg) has been inserted into the vaccinia virus genome under control of vaccinia virus early promoters. Cells infected with these vaccinia virus recombinants synthesize and excrete HBsAg and vaccinated rabbits rapidly produce antibodies to HBsAg.

IN 1798, Jenner described vaccination as a method of preventing smallpox¹. Inoculation with vaccinia virus provided a high degree of immunity against variola virus, the aetiological agent of smallpox, because of the live nature of the vaccine and the

close genetic relationship of the two viruses. Recently, similar uses of vaccinia virus recombinants that express genes of unrelated pathogens were proposed²⁻⁴. The feasibility of constructing such chimera was demonstrated by the molecular cloning

and expression of the herpesvirus thymidine kinase (*TK*) gene in vaccinia virus^{3,4}. We now describe vaccinia virus recombinants that synthesize hepatitis B virus surface antigen (HBsAg) on infection of cells and that induce the production of antibodies to HBsAg after inoculation of experimental animals.

The strategy previously used in our laboratory to clone and express foreign DNA in vaccinia virus took into account the large genome size and transcriptional regulatory signals used by the viral RNA polymerase⁴. A chimaeric gene consisting of a vaccinia virus promoter^{5,6} fused to a foreign protein coding sequence was flanked by vaccinia virus DNA within a plasmid vector. Transfection procedures^{7,8} were used to introduce the construct into vaccinia virus-infected cells where homologous recombination with genomic viral DNA occurred. Virus recombinants were selected either by their expression of herpesvirus *TK* or by loss of vaccinia virus *TK* expression. The latter was achieved by insertion of foreign DNA into the coding sequence of the previously mapped vaccinia virus *TK* gene^{9,10}. Significantly, expression of foreign DNA was dependent on vaccinia virus transcriptional regulatory sequences⁴.

To facilitate the use of vaccinia virus as a general expression vector, we have constructed plasmids (manuscript in preparation) that contain multiple restriction endonuclease sites introduced next to the vaccinia virus *TK* promoter¹⁰ or next to another early promoter⁵ translocated within the body of the *TK* gene. These plasmids enable any continuous protein coding sequence to be inserted next to a vaccinia promoter. Potential problems arising from improper codon phasing or unpredictable properties of fusion proteins are avoided by juxtaposing the transcriptional initiation site of the vaccinia promoter and the translational initiation codon of the foreign gene. As the plasmid origin of replication and the antibiotic resistance gene are retained, the construct can be selected and amplified in *Escherichia coli*. After transfection of wild-type vaccinia virus-infected cells with the chimaeric plasmid, *TK*⁻ recombinants are selected by plaque assay in human *TK*-143 cells in the presence of 5-bromodeoxyuridine (BUdR)⁴.

Construction of vaccinia virus recombinants containing the HBsAg gene

The hepatitis B viral genome has been cloned in phage λ and plasmid vectors and its primary structure has been determined¹¹⁻¹⁶. A continuous sequence that could encode surface antigen was identified and expression of this gene has been obtained in *E. coli*¹⁷ and yeast¹⁸ and by cells infected with SV40 recombinants^{19,20}. The 1.35-kilobase (kb) *Bam*HI fragment used to express HBsAg subtype adw in an SV40 vector¹⁹ was selected for the present experiments. Using protocols outlined above, the 1.35-kb *Bam*HI fragment was inserted into several different plasmid vectors linearized with *Bam*HI. Of the five plasmid recombinants formed, two (pHBs2 and pHBs4) have a vaccinia virus promoter and transcriptional initiation site translocated from an early gene encoding a 7,500 molecular weight (7.5K) polypeptide^{4,5} next to the beginning of the HBsAg gene, two controls (pHBs1 and pHBs3) have the translocated promoter adjacent to the opposite end of the HBsAg gene, and one (pHBs5) has the vaccinia virus *TK* promoter and transcriptional initiation site next to the beginning of the HBsAg gene (Fig. 1).

All the plasmids were analysed by restriction endonuclease digestions to confirm their proper construction and then used to transfect cells infected with wild-type vaccinia virus in P2 conditions as required by the NIH guidelines. Approximately 60% of *TK*⁻ plaques were shown to contain hepatitis B virus DNA by a rapid dot-blot hybridization procedure⁴. No cross-hybridization between DNAs from hepatitis B virus, vaccinia virus or uninfected cells was noted. A second plaque purification in BUdR was performed and all previously positive virus isolates again hybridized to hepatitis B virus DNA. Twice plaque-purified virus stocks were then amplified by passage in *TK*⁻

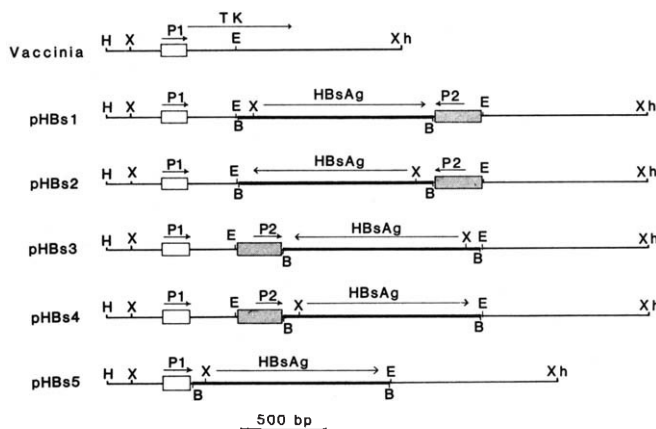


Fig. 1 Structures of chimaeric genes. A 2-kb segment from the left side of the *Hind*III J fragment of vaccinia virus is shown in the first line. The location and direction of the *TK* structural gene and putative promoter (P1) are indicated. In pHBs1 to pHBs4 (lines 2–5), chimaeric genes containing a translocated vaccinia virus promoter (P2) fused to a 1.3-kb hepatitis B virus *Bam*HI DNA fragment are inserted at the *Eco*RI site within the *TK* gene. The location and direction of the HBsAg structural gene and translocated promoter are indicated. In pHBs5, ~300 bp of the vaccinia virus *TK* gene has been deleted, multiple restriction endonuclease sites have been engineered next to the *TK* promoter, and the same hepatitis virus DNA fragment has been inserted. Details regarding the construction of the plasmid recombinants used for cloning the hepatitis B virus DNA fragment will be described elsewhere (manuscript in preparation). Abbreviations for restriction endonuclease cleavage sites are: H, *Hind*III; X, *Xba*I; E, *Eco*RI; Xh, *Xho*I; B, *Bam*HI.

Table 1 Production of HBsAg and vaccinia virus from infected cells

	ng HBsAg per 5×10^6 cells		Virus yield (PFU)	
	Cell extract	Culture medium	Cell extract	Culture medium
Uninfected	<1	<1	ND	ND
WT	<1	<1	7.8×10^8	7.7×10^7
vHBs1	11	20	8.3×10^8	10.2×10^7
vHBs2	835	1,700	7.9×10^8	9.9×10^7
vHBs3	14	25	9.1×10^8	9.0×10^7
vHBs4	930	1,700	8.8×10^8	9.8×10^7
vHBs5	35	80	10.3×10^8	9.6×10^7
Hepatoma cells	340	900	ND	ND

CV-1 cell monolayers were infected with purified wild-type (WT) or vaccinia virus recombinants (vHBs 1–5) at 30 plaque forming units (PFU) per cell or mock infected. At 2 h, virus inoculum was replaced with 2.5 ml of Eagle medium containing 2.5% fetal bovine serum. Cells were collected at 24 h and separated from culture medium by centrifugation at 2,000g for 5 min. Cell pellets were suspended in 2.5 ml of phosphate-buffered saline, freeze thawed three times and sonicated. Equal portions of cell extracts and culture medium were tested for HBsAg by radioimmunoassay and for vaccinia virus by plaque assay in CV-1 cells. Culture medium and a cell extract prepared as above from a hepatoma cell line PLC/PRF/5 3 days after confluency were tested in parallel for HBsAg.

cells with BUdR selection and then in CV-1 cells without selection.

Localization of hepatitis B virus DNA within the vaccinia virus genome

Recombinant virus DNA was examined by restriction endonuclease analysis and agarose gel electrophoresis. Results obtained with one pair of recombinants (vHBs1 and vHBs2) are shown in Fig. 2. Hybridization with ³²P-labelled vaccinia virus DNA (Fig. 2a) demonstrated that the 5.1-kb *Hind*III J fragment of wild-type (WT) virus was replaced in DNA of recombinant viruses by a larger one of 6.8 kb that migrated

Fig. 2 Autoradiographs of DNA blots containing restriction endonuclease fragments of plasmid or virus recombinants probed with ^{32}P -labelled vaccinia virus or hepatitis B virus DNA. DNA from wild-type virus (WT) or indicated plasmid or virus recombinant was digested with *Hind*III, *Bam*HI or *Xba*I restriction endonucleases. Fragments were separated by electrophoresis in a 0.6% agarose gel and transferred to two sheets of nitrocellulose placed on either side of the gel. Blots were baked at 80 °C for 2 h and then incubated for 4 h at 65 °C with 5 $\times$ Denhardt's solution²⁸ in 0.6 M NaCl, 0.06 M Na citrate (5 $\times$ SSC) containing 0.1 mg ml⁻¹ of denatured salmon sperm DNA. Either nick-translated denatured vaccinia virus DNA (a) or hepatitis B virus *Bam*HI fragment DNA (b) was added and the incubation was continued for 15 h. Nitrocellulose sheets were washed twice for 15 min at 65 °C with 2 $\times$ SSC containing 0.1% SDS and twice with 0.2 $\times$ SSC containing 0.1% SDS. The sizes of DNA fragments in kb are shown.

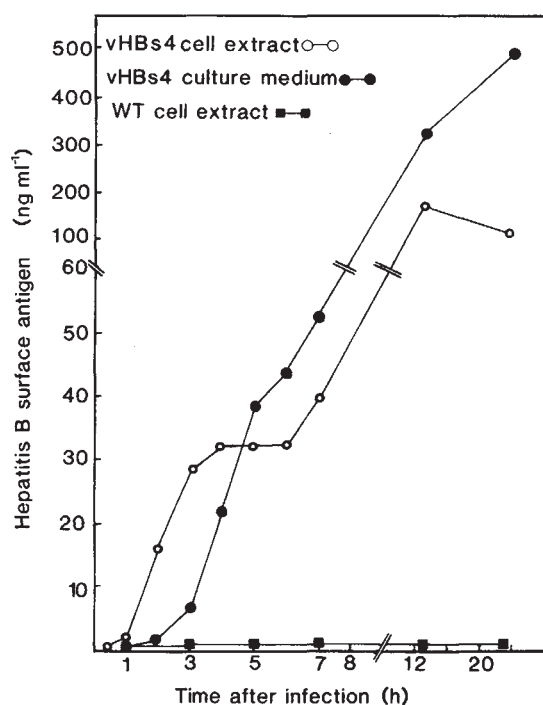
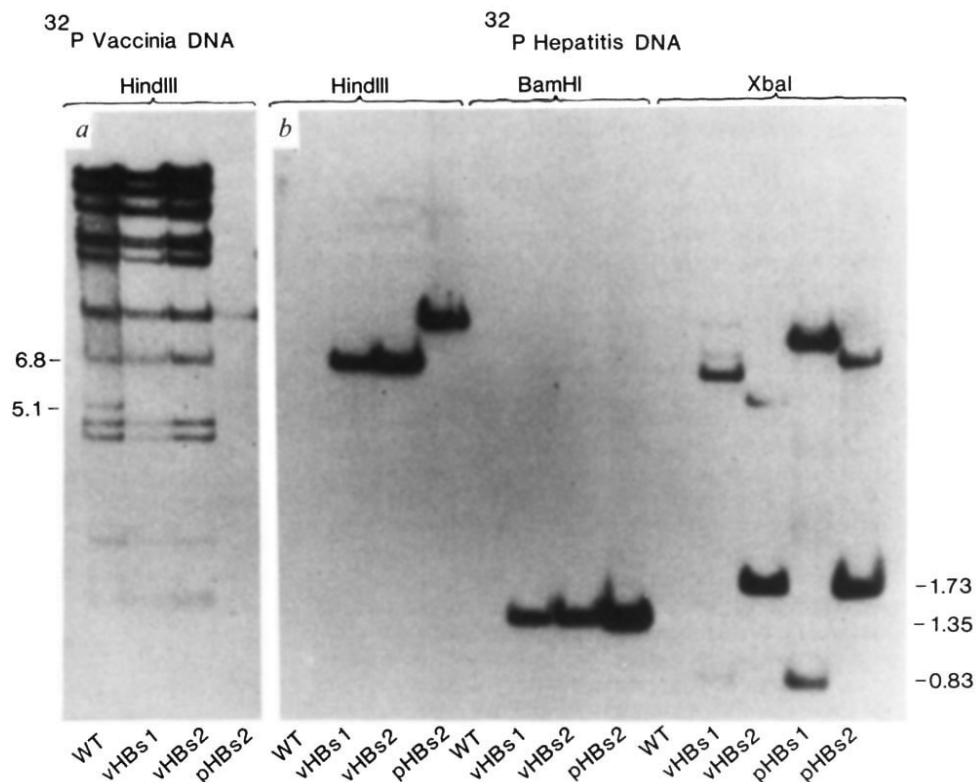


Fig. 3 Time course of synthesis and excretion of HBsAg. Monolayers (16 mm diameter) of CV-1 cells were infected with purified wild-type (WT) or vHBs4 at a multiplicity of 30 PFU per cell. After 1 h, the virus inoculum was replaced with 0.2 ml of Eagle medium containing 2.5% fetal bovine serum. At various times, the monolayers were collected and cells were separated from the medium by centrifugation for 10 s at 12,000g. Cell extracts were prepared by resuspending the cells in 0.2 ml phosphate-buffered saline, freezing and thawing three times, and sonicating. Equivalent samples of cell extract and medium were tested for HBsAg by radioimmunoassay.

almost coincidentally with the *Hind*III I fragment. The 6.8-kb *Hind*III fragment of recombinant viruses was distinguished from the *Hind*III I fragment by hybridization to a ^{32}P -labelled hepatitis B virus DNA probe (Fig. 2b). A band of 9.3 kb was detected by hybridization of either probe to *Hind*III-linearized pHBs2 (Fig. 2a, b). *Bam*HI excised the 1.35-kb hepatitis virus DNA fragment from both vaccinia virus recombinants and from the plasmid (Fig. 2b). *Xba*I digestions confirmed the relative orientations of the DNA segments predicted from Fig. 1; the virus recombinants vHBs1 and vHBs2 had hepatitis DNA in fragments of 0.83 and 1.73 kb, respectively, co-migrating with fragments from pHBs1 and pHBs2 (Fig. 2b). Larger *Xba*I restriction fragments of virus recombinants and corresponding plasmids did not co-migrate because in one case the hepatitis DNA was fused to only vaccinia sequences and in the other was fused to vaccinia and plasmid sequences. These data, and similar data for the other recombinants, indicated that the chimaeric genes were properly inserted into the *TK* gene within the *Hind*III J fragment of vaccinia virus DNA. Moreover, no other genomic rearrangements were detected.

Purity of the virus recombinants was demonstrated in several ways. The blots in Fig. 2a did not reveal any traces of a 5.1-kb *Hind*III J fragment indicative of contamination with wild-type virus or spontaneous *TK*⁻ mutants. The absence of *TK*⁺ virus was also indicated by identical plaque titres in *TK*⁻ cells in the presence and absence of BUdR. Furthermore, when a cell monolayer containing several hundred *TK*⁻ plaques was transferred to nitrocellulose by the procedure of Villareal and Berg²¹, all hybridized to ^{32}P -labelled hepatitis B virus DNA.

Expression of HBsAg gene

Preliminary evidence for HBsAg expression was obtained by incubating cell monolayers containing recombinant virus plaques with ^{125}I -labelled antibody against HBsAg. A radioimmunoassay (AUSRIA II, Abbott) provided more quantitative evidence for HBsAg synthesis. The amounts of HBsAg in cell extracts and supernatant medium, 24 h after infection with recombinant virus, are presented in Table 1. Approximately 2.6 μg of HBsAg was detected per 5×10^6 cells infected with vHBs2 or vHBs4 whereas smaller but still significant amounts

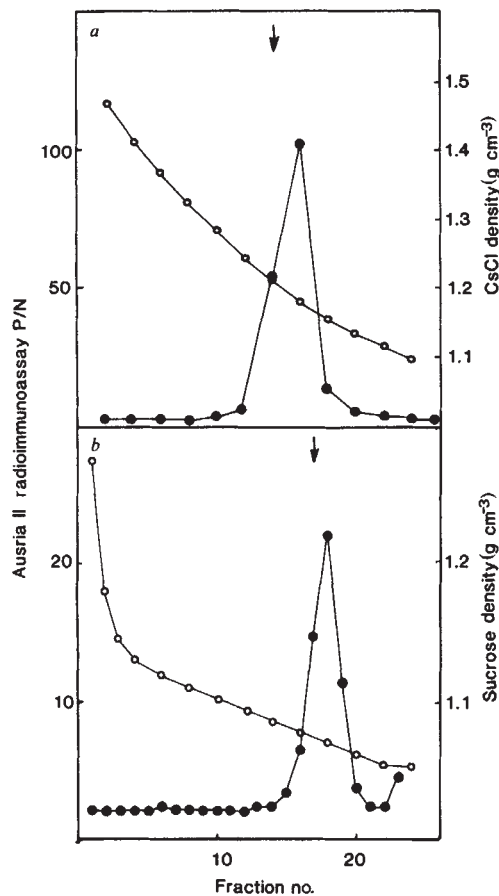


Fig. 4 Purification of HBsAg particles. *a*, Culture medium from CV-1 cells 16 h after infection with vHBs4 was clarified by low-speed centrifugation and then centrifuged at 75,000g for 24 h at 4°C in an SW41 rotor. Pelleted material was resuspended in 4.5 ml of 10 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA and then 1.38 g of CsCl was added to a final density of 1.2 g cm⁻³. Samples were centrifuged at 220,000g for 64 h in an SW65 Ti rotor at 4°C. Fractions (0.4 ml) were collected and diluted 1:20 for radioimmunoassay. Data are expressed as ratios of ¹²⁵I c.p.m. in sample (P) to mean negative control (N) of 103 c.p.m. CsCl density was determined by refractometry. *b*, Peak fractions from gradient similar to that in *a* were dialysed against phosphate-buffered saline, layered on a linear 10–30% (w/w) sucrose gradient in phosphate-buffered saline over a 0.5-ml cushion of 66% (w/w) sucrose, and centrifuged for 4.5 h at 150,000g in the SW41 rotor at 4°C. Fractions (0.5 ml) were collected, diluted 1:10 and radioimmunoassayed. Sucrose density was determined by refractometry. Arrows indicate the peak fractions of HBsAg particles purified in parallel from culture medium of hepatoma cell line PLC/PRF/5.

were made by vHBs5-infected cells. The levels of HBsAg detected in cells infected with vHBs1 and vHBs3 were only slightly higher than background values from uninfected or wild-type virus-infected cells. The very low levels of antigen synthesis in cells infected with vHBs3 are consistent with the inappropriate positions of both vaccinia virus promoters relative to the beginning of the HBsAg gene. Poor antigen synthesis by vHBs1 was also expected because the *TK* promoter precedes the HBsAg gene by about 500 nucleotides and the translocated promoter is downstream from the gene. The lower amounts of HBsAg synthesized by vHBs5 than by either vHBs2 or vHBs4 may reflect differences in activities of the *TK* and 7.5K gene promoters or the ways in which they were engineered. Collectively, these data indicate that HBsAg gene expression is under control of vaccinia virus promoters and that levels of antigen synthesis can exceed that produced by a hepatoma cell line (Table 1). Approximately two-thirds of the HBsAg was present in the culture medium 24 h after infection (Table 1). This

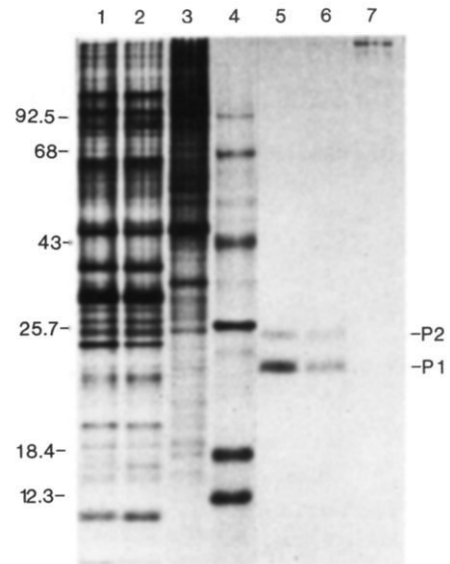


Fig. 5 Polyacrylamide gel electrophoresis of ³⁵S-methionine-labelled HBsAg. Monolayers of CV-1 cells were mock-infected or infected with purified vHBs2 or vHBs4 (30 PFU per cell) in Eagle medium containing 0.01 mM methionine. After 2 h, the virus inoculum was replaced with 0.5 ml of methionine-free medium supplemented with 120 μCi of ³⁵S-methionine (1,000 Ci mmol⁻¹, Amersham-Searle). The cells were incubated at 37°C for 4 h, washed three times with ice-cold phosphate-buffered saline, collected by centrifugation and lysed in 0.5 ml of 0.1 M Tris-HCl (pH 8.0), 0.1 M NaCl, 0.5% Nonidet P-40 detergent, 0.1% Aprotinin (Sigma). After 10 min at 0°C, the lysate was centrifuged and 80% of the supernatant was incubated with 25 μl of guinea pig non-immune serum at 4°C for 15 h. A 50-μl portion of a formalin-treated staphylococcal A suspension was added for 30 min at 4°C. After centrifugation, the supernatant was incubated for 4 h at 4°C with 20 μl of guinea pig HBsAg antiserum and then precipitated with staphylococcal A suspension. The pellet was washed twice with 0.05 M Tris-HCl (pH 7.5), 0.15 M NaCl, 0.1% SDS, 1% Triton X-100, 1% sodium deoxycholate and twice with 0.4 M LiCl, 2 M urea, 10 mM Tris-HCl (pH 8.0) and then solubilized with 50 μl of 0.06 M Tris-HCl (pH 6.8), 3% SDS, 5% β-mercaptoethanol, 10% glycerol, 0.002% bromophenol blue for 15 min at 25°C. After centrifugation, the supernatant was heated at 100°C for 1 min and applied to a 15% polyacrylamide gel. Samples containing 1% of the ³⁵S-methionine-labelled total cell extract were treated as above and analysed in parallel. After electrophoresis, the gel was fixed, treated with Enhance (NEN) and fluorographed. Lanes 1–3 contain total cell extracts from cells infected with vHBs4, vHBs2 or no virus, respectively. Lane 4 contains ¹⁴C-labelled marker proteins of indicated molecular weight. Lanes 5–7 contain immunoprecipitated polypeptides from cells infected with vHBs4, vHBs2 or no virus, respectively.

location was not a result of cell lysis because 90% of the infectious virus was still intracellular. In addition, the similar yields of wild-type and recombinant virus indicated that neither insertion of foreign DNA, the resulting *TK*⁻ phenotype, nor HBsAg expression impaired replication.

More detailed analysis of the synthesis and excretion of HBsAg in vaccinia virus recombinant infected cells is shown in Fig. 3. The detection of HBsAg in cell extracts as early as 1 h after infection was consistent with the linkage of the hepatitis gene to an early vaccinia virus promoter. By 2 h, HBsAg was detected in the medium and by 5 h, the amount in medium exceeded that associated with cells. At 23 h, approximately 80% of the HBsAg had been excreted.

Characterization of excreted HBsAg

HBsAg is present in the serum of hepatitis B virus carriers as particles with well defined biophysical properties²². When HBsAg excreted from cells infected with vaccinia virus

Table 2 Antibody response of vaccinated rabbits

Day	Anti-HBsAg (mIU ml ⁻¹)		
	Rabbit 1	Rabbit 2	Rabbit 3
1	—	0	2
2	—	0	2
3	1	0	2
4	—	17	6
5	—	4	6
6	5	10	4
7	—	66	26
8	—	174	259
9	1	242	365
10	—	212	253
11	—	139	180
14	2	111	106
17	—	41	133
20	1	33	581
24	—	53	1,965
31	0	34	2,563

Rabbits were injected intradermally with (1) 10^8 PFU of purified wild-type vaccinia virus at four sites, (2) 10^8 PFU of vHBs4 at four sites or (3) 10^8 PFU of vHBs4 at one site. Serum was prepared on the days indicated and assayed for anti-HBsAg by a radioimmunoassay (AUSAB, Abbott) using a positive human immunoglobulin (Bureau of Biologics, hepatitis B immune globulin, ref. no. 2)²⁹ as a standard. Background titres obtained with preimmune serum of each rabbit were subtracted. Typical skin lesions appeared on day 5 and healing began shortly thereafter.

recombinants was centrifuged in CsCl, radioimmunoassays revealed a peak at a density of 1.2 g cm^{-3} , similar to that obtained on simultaneous analysis of particles from the medium of hepatoma cell line PLC/PRF/5^{23,24} (Fig. 4a). The similarity of the HBsAg particles from the two sources was also shown by sucrose gradient sedimentation (Fig. 4b). 22 nm particles in the peak sucrose gradient fractions were visualized by electron microscopy.

Identification of HBsAg polypeptides

HBsAg consists of a major protein of apparent molecular weight 22,000–25,000 (P1) and a higher molecular weight glycoprotein (P2)^{24,25}. As seen in Fig. 5, two ³⁵S-methionine-labelled polypeptides with estimated molecular weights of 23,000 and 25,400 were specifically immunoprecipitated from cells infected with vHBs2 and vHBs4. The two polypeptides co-migrated with immunoprecipitated polypeptides from hepatoma cell line PLC/PRF/5.

Pulse-chase experiments more accurately defined the excretion time of HBsAg. After a 20-min pulse with ³⁵S-methionine and a chase with excess unlabelled amino acid, labelled immunoprecipitable HBsAg was found in the medium within 30 min and excretion continued for at least 4 h. The ratio of P2/P1 was greater in extracellular HBsAg than in the intracellular form, suggesting the glycosylation is closely coupled with excretion.

Vaccination of rabbits

To determine whether the vaccinia virus recombinants could elicit an appropriate antibody response in animals, rabbits were inoculated intradermally with vHBs4 or wild-type vaccinia virus. Following vaccination, the rabbits all developed typical local skin reactions but no HBsAg or vaccinia virus was detected in their serum. However, in the serum of both animals inoculated with vHBs4, anti-HBsAg titres of more than 200 mIU ml⁻¹ were detected within 9 days and in one of the animals the titre exceeded 2,500 mIU ml⁻¹ by day 31 (Table 2). Although not directly comparable, previous studies indicated that the minimal level of anti-HBsAg required for full protection of humans against hepatitis B virus infection was 10 mIU ml⁻¹ (ref. 26).

Conclusions

We have described here infectious vaccinia virus recombinants that express the HBsAg gene under control of vaccinia virus early promoters. Synthesis of HBsAg was detected within 1 h of infection of cells and continued for more than 13 h. By 2 h, HBsAg was found in the culture medium, by 5 h the level in the medium exceeded that associated with cells, and by 24 h it comprised 65–80% of the total. As 90% of the vaccinia virus produced was still cell associated, it is clear that HBsAg was specifically excreted. Approximately 2.6 µg of HBsAg was produced in 24 h by 5×10^6 cells in 2.5 ml of medium. By comparison, we found that the amounts of HBsAg accumulated by 5×10^6 hepatoma cells at 3 days after confluence was about 1.2 µg; in defective SV40 vector systems, values of 1.25 µg and 19 µg per 5×10^6 cells were reported^{19,20}.

In all respects examined (antigenicity, polypeptide composition, buoyant density and sedimentation rate), the HBsAg made by vaccinia virus recombinants was similar or identical to material obtained from the serum of human hepatitis B virus carriers. Moreover, rabbits vaccinated with purified recombinant vaccinia virus exhibited typical local skin lesions and produced antibodies to authentic HBsAg. In humans, the equivalent anti-HBsAg titres would provide protection against hepatitis B virus infection.

In view of the success of these experiments, it is pertinent to question whether the potential live vaccine described here would offer advantages over alternative purified subunit vaccines that already or soon will exist. In this context, it is important to consider that hepatitis B virus infection is a worldwide health problem²⁷. Approximately 200 million people are chronically infected and large numbers of deaths are attributed to fulminant hepatitis, cirrhosis and primary hepatocellular carcinoma. The inability to grow hepatitis B virus in tissue culture or in animals other than chimpanzees and the absence of naturally occurring attenuated strains prevented the development of a conventional live vaccine. However, the presence of large amounts of HBsAg in the blood of infected individuals provided a source for an effective subunit vaccine²⁷. Small amounts of HBsAg have also been made in recombinant bacteria¹⁷ and yeast¹⁸ and with further refinements this might provide a practical alternative source of purified HBsAg. Other methods involving the use of synthetic polypeptides are promising but uncertain. While multiple injections with HBsAg have already proved their value in protecting high risk groups in low endemic areas of Europe and North America, the high cost associated with purification, safety testing and use pose problems for their widespread application in high endemic areas of Africa and Asia where treatment of entire populations may be necessary.

Undoubtedly, the campaign to eradicate smallpox has been the most successful mass vaccination programme carried out in undeveloped countries. This achievement may be attributed to a number of factors, including the potency, stability and low cost of the live vaccine, as well as the ability of medically unskilled personnel to administer it in non-sterile conditions as a single dose on a mass scale. The efficiency of this method and the enormous experience with it have encouraged us to pursue the development of a vaccinia virus recombinant hepatitis B vaccine for immunization in areas where the risks of hepatitis greatly outweigh the low incidence of side effects of vaccination. Towards this goal, long-term experiments to evaluate the safety and efficacy of this vaccine in susceptible chimpanzees have been initiated. In addition, since stable and infectious vaccinia virus recombinants that contain 25,000 base pairs (bp) of foreign DNA have been constructed (G.L.S., unpublished), it may be possible to prepare polyvalent vaccines against a number of different pathogenic agents.

We thank B. H. Hoyer and J. L. Gerin for providing hepatitis B virus recombinant DNA, Research Resources Section of National Institute of Allergy and Infectious Diseases for HBsAg antiserum, J. Shih for human anti-HBsAg immunoglobulin, D. Djurickovic for vaccinating rabbits and collecting serum

samples, A. Weir and W. T. Ruyechan for electron microscopy, J. Janik for advice regarding immunoprecipitations, N. Cooper for technical assistance, and J. Carolan for typing the manuscript.

Received 25 November 1982; accepted 31 January 1983.

1. Jenner, E. *An Inquiry into the Causes and Effects of the Variolae Vaccinae, a Disease Discovered in some Western Counties of England, particularly Gloucestershire, and known by the Name of Cow Pox*. 1798. (Reprint: Cassell, London, 1896).
2. Weir, J. P., Bajszar, G. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1210–1214 (1982).
3. Panicali, D. & Paoletti, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4927–4931 (1982).
4. Mackett, M., Smith, G. L. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7415–7419 (1982).
5. Venkatesan, S., Baroudy, B. M. & Moss, B. *Cell* **125**, 805–813 (1981).
6. Venkatesan, S., Gershowitz, A. & Moss, B. *J. Virol.* **44**, 637–646 (1982).
7. Sam, C. K. & Dumbell, K. R. *Ann. Virol. Inst. Pasteur, Paris* **132E**, 135–150 (1981).
8. Nakano, E., Panicali, D. & Paoletti, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1593–1596 (1982).
9. Hruby, E. D. & Ball, L. A. *J. Virol.* **43**, 403–409 (1982).
10. Bajszar, G., Wittek, R., Weir, J. P. & Moss, B. *J. Virol.* **45**, 62–72 (1983).
11. Charnay, P., Pourcel, C., Louise, A., Fritsch, A. & Tiollais, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2222–2226 (1979).
12. Burrell, C. J., Mackay, P., Greenway, P. J., Hofschneider, P. H. & Murray, K. *Nature* **279**, 43–47 (1979).
13. Sninsky, J. J., Siddiqui, A., Robinson, W. S. & Cohen, S. N. *Nature* **279**, 346–348 (1979).
14. Valenzuela, P. *et al. Nature* **280**, 815–819 (1979).

Note added in proof: Analysis of sera for 94 days post-vaccination indicated that the anti-HBsAg titres increased, reaching values of 540 and 9,600 mIU ml⁻¹ for rabbits 2 and 3, respectively.

15. Galibert, F., Mandart, E., Fitoussi, F., Tiollais, P. & Charnay, P. *Nature* **281**, 646–650 (1979).
16. Pasek, M. *et al. Nature* **282**, 575–579 (1979).
17. MacKay, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 4510–4514 (1981).
18. Valenzuela, P., Medina, A., Rutter, W. J., Ammerer, G. & Hall, B. D. *Nature* **298**, 347–350 (1982).
19. Moriarty, A. M., Hoyer, B. H., Shih, J. W.-k., Gerin, J. L. & Hamer, D. H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2606–2610 (1981).
20. Liu, C.-C., Yansura, D. & Levinson, A. D. *DNA* **1**, 213–221 (1982).
21. Villarreal, L. P. & Berg, P. *Science* **196**, 183–185 (1977).
22. Gerin, J. L., Holland, P. V. & Purcell, R. H. *J. Virol.* **7**, 569–576 (1977).
23. Macnab, G. M., Alexander, J. J., Letcasas, G., Bey, E. M. & Urbanowicz, J. M. *Br. J. Cancer* **34**, 509–515 (1976).
24. Marion, P. L., Salazar, F. H., Alexander, J. J. & Robinson, W. J. *Virol.* **32**, 796–802 (1979).
25. Peterson, D. L. *J. biol. Chem.* **256**, 6975–6983 (1981).
26. Courouce-Pauty, A. M., Naret, C., Cianconni, C., Adhemar, J. P. & Soulier, J. P. *Transplant. clin. Immun.* **10**, 77–85 (1978).
27. Purcell, R. H. & Gerin, J. L. *Rev. infect. Dis.* **2**, 470–492 (1980).
28. Denhardt, D. T. *Biochem. biophys. Res. Commun.* **23**, 641–646 (1966).
29. Gerety, R. J., Smallwood, L. A. & Tabor, E. *New Engl. J. Med.* **303**, 529 (1980).

Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations

Charlotte Paquin* & Julian Adams†

Division of Biological Sciences, University of Michigan, Ann Arbor, Michigan 48109, USA

Adaptive mutations are shown to have a higher frequency of fixation in evolving diploid than in haploid populations of the yeast Saccharomyces cerevisiae, providing direct evidence that it may be an evolutionary advantage to be diploid.

A DOMINANT diploid phase is one of the major features of the life cycle of almost all eukaryotes, and can be considered to be circumstantial evidence for the adaptive significance of this phase. However, this subject has received scant attention from evolutionary biologists, especially when compared with the extensive literature on the adaptive significance of sex or recombination, a closely related, but independent issue (see refs 1, 2 and references therein). As early as 1929, Svedelius³ argued that adaptive significance of diploidy was due to the ability of diploids to generate more variability than haploids. This argument and similar ones advanced by others^{2,4–6} can be considered as belonging to the balance school⁷ of population genetics, which asserts that genetic variation in natural populations is (1) ubiquitous and (2) somehow adaptive. A contrasting theory, which really owes its allegiance to the neoclassical school of population genetics, argues that the adaptive significance of the diploid phase lies in its ability to mask deleterious recessive mutations in either somatic^{5,8,9} or germ-line cells^{5,10–12}.

There exist significant criticisms of both sets of theories. On the one hand, many of the postulated advantages of diploidy advanced by the balance school are unlikely to have been manifested as soon as an organism converted to diploidy and may only be consequences rather than causes of a dominant diploid phase. Furthermore, the adaptive significance of variation must involve either heterozygous advantage, considered by some^{7,13,14} to be an uncommon phenomenon in natural populations, or mechanisms of group selection which are by no means universally accepted. On the other hand, the advantages for diploidy postulated by the neoclassical school will be on the order of the genomic mutation rate. These values are extremely small for populations of microorganisms (1×10^{-3}), although they may be appreciable for more complex organisms such as

*Drosophila*¹⁵. Nevertheless, the greater accumulation of deleterious alleles in diploids will result in a higher mutational load as equilibrium is approached, resulting in an adaptive disadvantage for diploidy.

Here we investigate an alternative hypothesis for the adaptive significance of the diploid phase—namely that the rate of adaptation of diploid cells to new environments may be faster than that of haploid cells. The rationale for this hypothesis is simple; diploids can be considered to have undergone a duplication for every gene in the genome. As such, the rate of occurrence of adaptive mutations in diploids should be twice that of haploids, and so they should enjoy increased rates of adaptation. Similar hypotheses have been advanced to explain the success of organisms that have undergone gene duplications^{16–20}, and of polyploids²¹. Diploidy may be considered to be no more than a special case of polyploidization.

An important assumption implicit in this hypothesis is that most (>50%) adaptive mutations are expressed in the heterozygous state; *a priori* this assumption is not unreasonable. Advantageous mutations may be expected to improve or change gene function, and hence be dominant or semi-dominant whereas deleterious mutations are likely to cause the loss or attenuation of function and thus be recessive. The available data are not particularly revealing on this issue. Most mutations identified in *Drosophila* are recessive²² whereas most of those identified in humans are dominant²³. However, these results are difficult to interpret as few advantageous mutations have been studied; the mutations studied may be a biased sample of all mutations, and the results may vary from organism to organism.

Rates of adaptation to a new environment, in this case a carbon-limited chemostat environment, were measured in *Saccharomyces cerevisiae*. This organism is particularly suitable for such studies as it is a single-cell eukaryote with both stable haploid and diploid phases. Furthermore, it has been well characterized both physiologically and genetically.

* Present address: ARCO Plant Cell Research Institute, 6560 Trinity Court, Dublin, California 94566, USA.

† To whom reprint requests should be addressed.

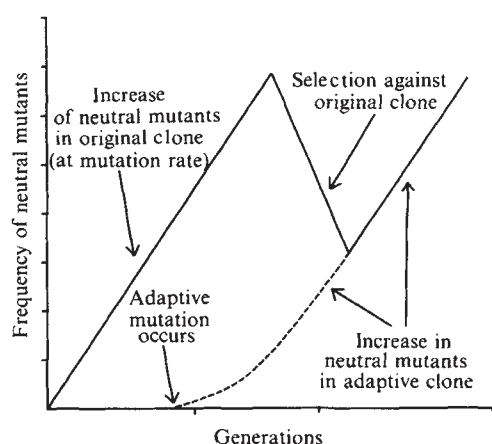


Fig. 1 Diagrammatic representation of the fluctuation in the frequency of neutral mutants in an asexual population showing the occurrence of, and selection for, an adaptive mutation.

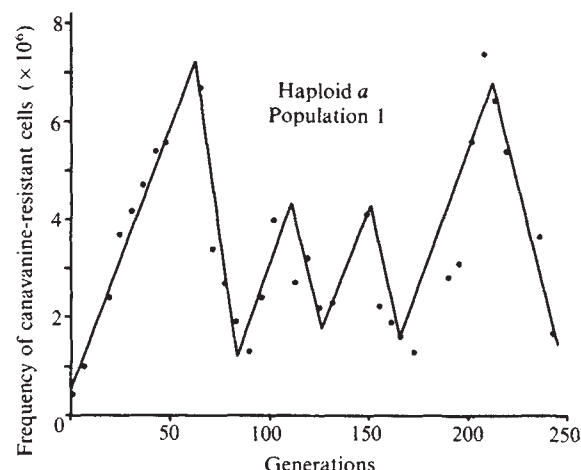


Fig. 2 Fluctuation in the frequency of canavanine resistance in one population of the haploid *a* strain, CP1AB-1A²⁸. The line illustrates the fluctuations in the frequency of canavanine resistance expected assuming the occurrence and fixation of four adaptive mutations. The cells were grown in minimal medium + 0.08% glucose⁶ in continuous culture with a dilution rate of $\sim 0.20 \text{ h}^{-1}$. Increasing the concentration of glucose in the medium above 0.08% resulted in a proportionate increase in the population size, demonstrating that glucose was the limiting substrate. Selection for mutants such as *abs* and *wal*⁵⁶, which cause cells to clump and sink to the bottom of continuous culture apparatuses, was reduced by pumping medium and cells from the bottom of the continuous culture apparatuses at $\sim 67\%$ of the dilution rate. Despite this, clumping cells were selected for after 100–300 generations. Each chemostat was sampled about every 24 h. Cells were sonicated for 15–30 s (Branson sonicator Model 185 with a micro-tip; power setting 3; equivalent to $\sim 15 \text{ W}$) to break up clumps of cells, then plated onto minimal medium at appropriate dilutions to determine the number of viable cells per ml, and plated onto minimal medium plus canavanine⁶ to determine the frequency of canavanine resistance. Canavanine resistance is recessive, determined by a single locus (*can1*)⁵⁷, and is neutral in our experimental conditions (ref. 6 and Table 1). An unsonicated aliquot was frozen in 15% glycerol at -70°C for later analysis. Experiments were stopped when clumped cells could not be separated by sonication. The frequency of canavanine resistance was estimated as the number of mutant cells per ml divided by the number of cells per ml as determined from the viable counts. The population size for each sample was estimated as the number of cells per ml as determined by the viable count multiplied by the volume of the chemostat. The number of generations was determined from $t \times \ln 2/D$ where D is the dilution rate (h^{-1}) and t is the time (h) from the start of the the experiment⁵⁸. The number of cell generations between samples was determined as the average of the population sizes for each sample weighted by the number of generations between samples. At the end of the experiment, the cells were tested for mating ability, sporulation, ploidy and for complementation with *gal2* and *gal3* tester strains²⁸ to verify that the cells had the same genotype as the strain used to inoculate the chemostat.

Table 1 Relative fitnesses of the adaptive clones isolated from the haploid *a* population (see Fig. 2)

Strain 1		Strain 2		Fitness of strain 2 relative to a fitness of 1 for strain 1 $\pm$ s.e.m. per generation
Isolated at generation	Putative no. of adaptive mutations	Isolated at generation	Putative no. of adaptive mutations	
0	0	0*	0	0.99 ± 0.01
0*	0	30	0	1.02 ± 0.01
0.5	0	48*	0	1.02 ± 0.01
25*	0	97	1	1.10 ± 0.01
30	0	97*	1	1.11 ± 0.01
97*	1	133	2	1.05 ± 0.01
133*	2	203	3	1.16 ± 0.02
203*	3	245	4	1.10 ± 0.01

For each experiment the strain which had been isolated after the smaller number of generations of growth in the chemostat is designated as strain 1. * Indicates that the strain was marked with canavanine resistance by plating $\sim 1 \times 10^8$ cells on to minimal medium + 2 mM canavanine sulphate and picking the spontaneous mutant colonies that appeared. At least five colonies of each of the two strains to be compared, were picked, pooled and inoculated into 50–75 ml of minimal medium and grown overnight at 30°C with shaking in a water bath. The cells were then inoculated into a chemostat in approximately equal frequency at approximately the cell densities observed in long-term experiments $5 \times 10^7 \text{ cells ml}^{-1}$. Medium flow commenced immediately after inoculation. Samples were taken every 12–24 h for 3–5 days. The samples were sonicated as described in Fig. 2 legend, plated onto minimal medium after appropriate dilution, incubated at 30°C for 3–5 days and then replica-plated onto selective media to monitor changes in the frequency of the resistance marker. Approximately 500 colonies were counted for each sample. Fitnesses ($\pm$ s.e.m.) were calculated as described previously⁵², relative to a fitness value of unity for strain 1.

Detection of adaptive changes

It is of course necessary that all adaptive mutations be detected to obtain unbiased estimates of the rate of adaptation in haploid and diploid yeast populations. In a sexually reproducing population this would be unreasonably difficult as it would require (1) *a priori* knowledge of the phenotypes of all possible adaptive mutations and (2) that all adaptive mutations be unambiguously detectable. However, in an asexually reproducing population the spread of a selectively favoured gene throughout the population may be detected without a knowledge of the phenotype of the adaptive mutation by monitoring the frequency of an unrelated neutral mutation. In such populations recombination

is virtually absent so all genes may be considered to be completely linked. In a large population (that is, 10^8 – 10^{10} organisms) in which the initial frequency of the neutral mutation is very close to zero, it will increase in frequency at the mutation rate such that after t generations, the frequency will be about μt (ignoring reverse mutation), where μ is the mutation rate. If at some point an adaptive mutation occurs in the population, the new adaptive clone will increase in frequency and both the old clone and the neutral mutants it has accumulated will be selected against. The probability that the adaptive mutation will occur in a cell carrying the neutral mutation will equal the frequency of the neutral mutant, $\sim \mu t$, which for small t , will be extremely small, of the order of 1×10^{-5} to 1×10^{-7} . As the

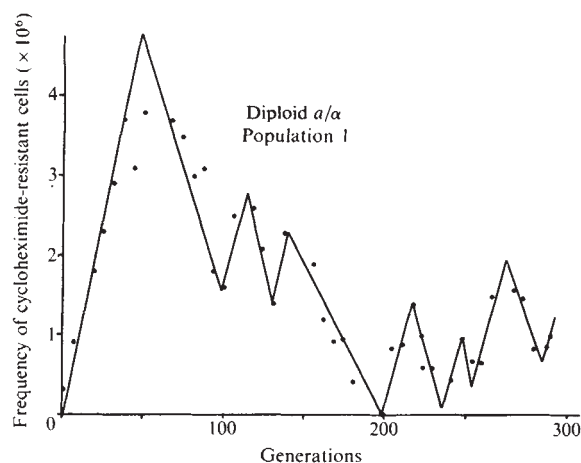


Fig. 3 Fluctuation in the frequency of cycloheximide resistance in one population of the diploid a/α strain, CP1AB-1AB²⁸. The line illustrates fluctuations in the frequency of cycloheximide resistance expected assuming the occurrence and fixation of six adaptive mutations. The experiment was carried out using the procedures described in Fig. 2 legend with the following modifications. The frequency of dominant cycloheximide resistance, determined by a single locus (*CHY4*)^{55,57}, was determined by plating cells onto minimal medium + 1.8 μ M cycloheximide. Dominant cycloheximide resistance is selected against in our experimental conditions (selective coefficient = 0.10). The genotype of the cells at the end of the experiment was determined by sporulating the cells and assaying the spores as described in Fig. 2 legend.

new adaptive clone has existed for fewer generations than the old clone it will have accumulated fewer neutral mutants and consequently there will be a net decrease in the frequency of neutral mutants. As cells from the old clone become rare the frequency of neutral mutants will again rise as the new clone accumulates mutants at the mutation rate, until the next adaptive mutation appears and the cycle repeats. A diagram of this cycle is shown in Fig. 1. Slightly deleterious mutations may also be used to detect adaptive changes in populations. They will show the same cyclic fluctuations in frequency except that the rate of increase of a slightly deleterious mutant will be less than the mutation rate by an amount proportional to the selective coefficient. In general this will not compromise their utility in detecting adaptive change in populations, provided the selective coefficients are small²⁴. These fluctuations are known in the microbiological literature as periodic selection²⁵. If there is tight linkage between the adaptive mutation and the neutral mutation monitored, similar correlated responses in gene frequency change may also occur in sexual populations, and in such cases the phenomenon has been termed 'hitchhiking'²⁶. Thus, by following the frequency of neutral or slightly deleterious mutants over time in an asexual population, the number and frequency of adaptive changes may be determined without a knowledge of the phenotype of the adaptive mutant itself. Furthermore, the possibility exists of comparing the rate of fixation of two populations using this approach²⁷.

Analysis of fluctuations in mutant frequencies

In the early experiments, fluctuations in the frequencies of three different mutant phenotypes were monitored in asexual haploid and diploid populations of *S. cerevisiae* grown in glucose-limited chemostats. The frequency of canavanine resistance, cycloheximide resistance and 5-fluorouracil resistance was monitored in the haploid populations and cycloheximide and 5-fluorouracil resistance in the diploid populations. Canavanine resistance is recessive and so could not be monitored in the diploid populations. All markers monitored in these populations showed the same pattern of fluctuations, confirming that the fluctuations observed were not unique properties of the

particular resistance marker. Therefore, in further experiments only canavanine resistance was monitored in the haploid populations, and cycloheximide resistance alone was monitored in the diploid populations as these two markers proved to be the easiest to score.

Figures 2 and 3 show the change in the frequency of canavanine- and cycloheximide-resistant mutants versus generation time in two representative populations, a haploid a population and a diploid a/α population, respectively. These two representative populations were analysed in detail to verify that the fluctuations in canavanine resistance (haploid population) and cycloheximide resistance (diploid population) were caused by the fixation of adaptive mutations in the populations.

The putative adaptive changes were reconstructed by isolating cells at different generations, and placing them in direct competition with each other in the same glucose-limited chemostat environment, one of the strains being marked with canavanine resistance in the case of the haploids and 5-fluorouracil resistance in the case of the diploids. The results (Tables 1, 2) are presented in the form of relative fitnesses—the convention in population genetics. Two sets of controls were performed for both haploids and diploids. In the first set (row 1 of Tables 1 and 2), the selective effect of the drug resistance marker was measured in the strains CP1AB-1A (haploid) and CP1AB (diploid), which had not been grown previously in glucose-limited chemostats. The results show that canavanine resistance is selectively neutral in these conditions, but that the 5-fluorouracil marker is selected against. Accordingly, most of the reconstruction experiments with diploids were carried out in duplicate, marking strain 1 with 5-fluorouracil resistance in one experiment, and strain 2 with 5-fluorouracil resistance in the other. Relative fitnesses were then calculated correcting for the selective effect of 5-fluorouracil.

The second set of controls (rows 2 and 3 of Tables 1 and 2) compared cells isolated at different generations, but which were still considered to be members of the same clone, as determined by the dynamics of the drug resistance marker frequency. Thus, in Table 1, cells isolated at the beginning of the experiment, at generation 0.5, were grown in competition with cells isolated at generation 48. The constant increase in the frequency of canavanine resistance until approximately generation 60 indicates that no adaptive change in the population should have occurred during that time. The results shown in Table 1 confirm the selective equivalence of cells isolated at these two times. A similar set of results (Table 2) was obtained for diploids.

Tables 1 and 2 also show the results of competition experiments between cells isolated before and after each adaptive shift. In every case fluctuations in the frequency of canavanine or cycloheximide resistance signalled the fixation of an adaptive mutation. In the case of the diploid population the reciprocal experiments showed no evidence of epistatic interactions between the gene for 5-fluorouracil resistance and the adaptive mutations. Surprisingly, the average fitness increment between the adaptive clones for the haploid population (0.10) was not significantly different from that (0.09) for the diploid population, suggesting that the adaptive mutations fixed in the haploid and diploid populations are equally advantageous.

It is also possible to calculate the selective coefficients directly from the sequential rates of decrease of the frequency of the neutral mutants in the long-term continuous culture experiments shown in Figs 2 and 3. The reliability of these estimates is questionable as they are based on few data, and also assume incorrectly that all mutants (canavanine- or cycloheximide-resistant) are derived from the strain not possessing the most recent adaptive mutation. Nevertheless these estimates (not shown) agree closely with the estimates from the reconstruction experiments.

Finally, to verify that the sequential increases in the frequency of the neutral mutants were consistent with the rates of mutation to canavanine and cycloheximide resistance respectively, these mutation rates were estimated independently by fluctuation tests (Table 3). These rates show an excellent correspondence

Table 2 Relative fitnesses of the adaptive clones isolated from the diploid a/α population (see Fig. 3)

Strain 1		Strain 2		Fitness of strain 2 relative to value of 1 for strain 1 ($\pm$ s.e.m.) per generation	Relative fitness of strain 2 corrected for selection of fluorouracil resistance marker
Isolated at generation	Putative no. of adaptive mutations	Isolated at generation	Putative no. of adaptive mutations		
0*	0	0	0	1.10 $\pm$ 0.01	—
0.5*	0	32	0	1.10 $\pm$ 0.01	1.00
0.5	0	44*	0	0.90 $\pm$ 0.01	1.01
26*	0	107	1	1.41 $\pm$ 0.03	1.25
26	0	107*	1	1.06 $\pm$ 0.02	1.16
101*	1	139	2	1.18 $\pm$ 0.04	1.06
101	1	139*	2	0.95 $\pm$ 0.06	1.04
139	2	213*	3	1.01 $\pm$ 0.03	1.10
213*	3	243	4	1.20 $\pm$ 0.02	1.09
213	3	243*	4	1.06 $\pm$ 0.03	1.16
243	4	268*	5	0.96 $\pm$ 0.03	1.05
262*	5	305	6	1.13 $\pm$ 0.04	1.03

Experiments were carried out as described in Table 1 legend with the following modifications. * Indicates that in each experiment, the strain was marked with 5-fluorouracil resistance. The selective medium was minimal medium + 0.2 mM 5-fluorouracil. This marker proved easier to score than cycloheximide resistance on the replica plates. Cells were inoculated at a cell density of 3×10^7 cells ml⁻¹, the approximate equilibrium cell density for diploids. Relative fitnesses ($\pm$ s.e.m.) were calculated as before (see Fig. 2 legend). Fitnesses of strain 2 were corrected for selection against the 5-fluorouracil resistance marker by assuming selective effects to be additive before normalizing the fitnesses relative to a fitness value of unity for strain 1.

with those estimated from the continuous culture experiments. In total, the independent estimates of the selective coefficients and mutation rates provide overwhelming support for the hypothesis that the fluctuations in the frequency of the resistant mutants, are a manifestation of the sequential occurrence and fixation of adaptive mutations.

Comparison of haploid and diploid populations

Rates of adaptation were measured for two haploid (a, α) and two diploid (a/α and a/a) strains which were isogenic²⁸ except for the mating type locus; 64 putative adaptive mutations were observed in 2,612 generations of 11 different populations. The results are summarized in Fig. 4. The average frequency of adaptive mutations fixed was 5.68×10^{-12} per cell generation for all diploids, whereas the average number of adaptive mutations fixed per cell generation for all haploid populations was 3.55×10^{-12} . The Wilcoxon Mann-Whitney non-parametric test (for example, see ref. 29) showed a highly significant difference ($P = 0.0002$) between the haploid and diploid populations. To test for an effect of the mating type locus, the same test was applied to the two sets (a and α) of haploid populations and the two sets (a/a and a/α) of diploid populations. As expected there was no significant difference in either case. Thus, the data show that diploids fix ~ 1.6 times as many adaptive mutations per cell as haploids. Furthermore, it can be seen that the rate of fixation of adaptive mutations does not decrease with time, as might be expected.

Frequency and nature of adaptive change

Fluctuations in the frequency of neutral mutants, signalling repeated adaptive genetic changes in *Escherichia coli*, were reported as early as 1950^{25,30}. Our results demonstrate for the first time that similar changes may occur in populations of the single-celled eukaryote *S. cerevisiae*, and show that the rate of adaptive change in different populations may be compared. Two aspects of these results are noteworthy.

(1) The genetic changes in both haploid and diploid populations occur surprisingly frequently, higher than has been reported previously. The only study that allows a direct comparison of the rate of adaptive change is that of Novick and Szilard^{30,31} concerning adaptation to a tryptophan-limited environment. The estimated frequency of fixation of adaptive mutations observed by these authors was $1.9\text{--}2.3 \times 10^{-12}$, values

significantly lower than those observed here. Clearly this difference may be explained by the different genetic and physiological characteristics of the two species but may also be explained by the difference in environmental conditions used. Novick and Szilard identified genetic changes occurring during adaptation to a tryptophan-limited environment where the number of genetic events leading to increased adaptation may be limited.

The identification of rapid adaptive genetic changes reported here has important implications for industrial applications of *S. cerevisiae* and related yeasts. Although some evidence exists for the occurrence of significant genetic changes during long-term continuous fermentation of the brewery yeast *Saccharomyces carlsbergensis*^{32,33}, no information has previously been available on their frequency. Genetic changes in the frequency reported here may seriously attenuate the economic properties of the organism. This problem may be particularly acute in industrial uses of *S. cerevisiae* as a host for artificial chimaeric plasmids (for example, see ref. 34), which themselves may be subject to frequent genetic changes³⁵ and high rates of segregational loss³⁶.

(2) Not only are the rates of adaptive change high in comparison with previous studies, but these rates of fixation of adaptive mutations did not decrease with the number of generations spent in the chemostat. Thus there is no evidence for the exhaustion of capacity for adaptive genetic change during the time period over which the populations were observed (up to 327 generations). The most obvious conclusion is that the number of loci and/or alleles for adaptive genetic change is extremely large. However, it is also possible that the epistatic or functional relationships between the adaptive mutations is sufficiently complex that adaptive mutations may cycle through the population over long time periods^{31,37}. This 'circular evolution' may lead to an infinitely repeating succession of adaptive changes with a limited repertoire of adaptive mutations. We present elsewhere³⁸ evidence from estimates of selective differences between clones differing in more than one adaptive mutation which suggests that such a phenomenon may be occurring in our populations. It has also been suggested to occur in evolving populations of RNA viruses³⁹.

Implications for the evolution of diploidy

The fundamental result obtained from these experiments is that the rate of adaptive change in diploid populations of

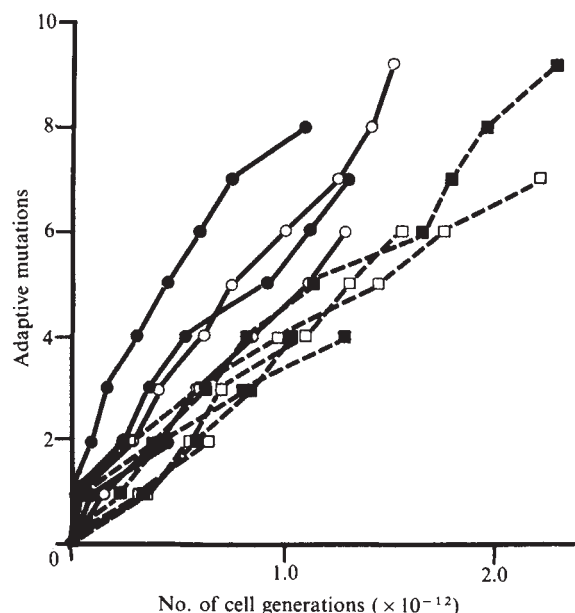


Fig. 4 Number of adaptive mutations fixed versus the number of cell generations elapsed for the 11 populations studied. Isogenic strains used to inoculate the continuous cultures were: CP1AB-1A haploid, *a* (■); CP1AB-1B haploid, α (□); CP1AB-1AA diploid, *a/a* (●); CP1AB diploid, *a/α* (○). The isolation and complete genotypes of these strains have been described previously²⁸. The mutation rates estimated for haploid *a* population 1 and diploid *a/α* population 1 are consistent with the data from the other nine populations and were used as the rate of increase in the resistant mutants in analysing the data from these populations. The number of adaptive mutations fixed in a population was determined by counting the number of decreases observed in the frequency of the cycloheximide- and canavanine-resistant mutants. Decreases in resistant mutant frequency were counted only when more than one sample showed a decrease and when it was consistent with our estimates of the mutation rates for such a decrease to have occurred. The time of occurrence of each adaptive mutation was assumed to be the time at which the first decrease in the frequency of the resistant mutant was observed. Although the actual time of occurrence of the adaptive mutations will be slightly earlier (see Fig. 1), this will not affect the estimates of the rates of adaptation because they depend on the elapsed time between adaptive mutations rather than their exact time of occurrence.

S. cerevisiae is 1.6 times higher than the rate in isogenic haploid populations. To our knowledge, this represents the first direct experimental demonstration of an adaptive advantage for the diploid phase of the life cycle of a eukaryote. The deviation of this value from the theoretically expected value of two, which assumes that all adaptive mutations are dominant, suggests that a small proportion are not expressed in the heterozygous state. However, the equality of the fitness increments between the adaptive clones in the haploid and diploid populations suggests that, of those adaptive mutations fixed in the diploid population, all exhibit complete dominance. This assumes, of course, that the spectra of adaptive mutations fixed in the two populations are the same. Clearly it would be of interest to compare the characteristics of the adaptive mutations and to measure directly their level of dominance. However it must be remembered that our approach for the detection of adaptive change does not allow the identification of the loci of the adaptive mutations. Without this information it is impractical to separate the combinations of adaptive mutations necessary for the direct measurement of dominance and the determination of the phenotypes of the individual mutations.

A higher rate of adaptive mutation for diploid cells should be common to all eukaryotes having an asexual diploid stage in the life cycle. Thus, it is tempting to suggest this as a general

Table 3 Mutation rate estimates

Mutant	Mating type	Estimate from continuous culture experiment	Estimate from fluctuation test	Published estimates
Canavanine resistance	<i>a</i>	1.06×10^{-7}	$9.7 \pm 5.0 \times 10^{-8}$	1.5×10^{-7} *
Cycloheximide resistance	<i>a/α</i>	8.66×10^{-8}	$2.0 \pm 1.5 \times 10^{-7}$	

The fluctuation test was carried out as described previously⁵³. Cells were grown in minimal medium and viable counts used to determine the total cell number. To estimate the mutation rate to cycloheximide resistance, cells were grown to a density of 5×10^5 cells ml⁻¹ to minimize background growth which commonly occurs at higher cell densities. More cultures (90 instead of 15) were grown and larger volumes (1 ml instead of 0.2 ml) were plated using a soft agar plating technique to offset the smaller number of cells per culture. Mutation rates and 95% confidence intervals for these rates were calculated using either the P_0 or median method of Lea and Coulson⁵⁴. Mutation rate estimates from the continuous culture experiments were the least-squares estimates of the slope of the initial increase in frequency of canavanine- or cycloheximide-resistant mutants versus the generation time. * Ref. 59.

mechanism for the evolution of diploidy. Although the rate of mutation may not be the limiting factor determining macroevolutionary change², this parameter may be critical in determining short-term adaptation to new environments for asexual organisms which cannot rely on recombination to generate variation in fitness. A similar argument has been used to explain the predominance of RNA viruses compared to DNA viruses³⁹, and at the microevolutionary level, to explain the selective advantage of the mutator alleles at the *mutT* and *mutH* loci in laboratory continuous cultures of *E. coli*⁴⁰⁻⁴³. In addition, on a theoretical level several workers⁴⁴⁻⁴⁶ have argued that in asexual populations selection may favour genotypes which increase mutation rate.

Although we consider a higher rate of adaptive change in diploids to be an important mechanism in the evolution of diploidy in many organisms, it is unlikely to explain satisfactorily the predominance of the diploid phase in all organisms. For organisms having no asexual diploid stage in the life cycle, such an effect may explain the maintenance of a predominant diploid phase, but not its evolutionary origin. In addition, Darlington⁴⁷ has suggested that a predominant diploid phase evolved independently at least six times, and several independent origins have been proposed for the ascomycetes alone⁴⁸. Indeed, it would be surprising if the mechanism of evolution of diploidy were the same in all cases. Clearly there is a link between multicellularity and a predominant diploid phase. Lewis and Wolpert⁴⁹ have suggested that this link may be causal and that diploids are more able to evolve the complex systems of gene regulation that are characteristic of multicellular organisms having complex somas. Protection from the deleterious effects of somatic mutation may also be more important in such organisms.

A further consideration is that the rate of fixation of adaptive mutations will also depend on population size. A higher frequency of adaptive mutation in diploids may be offset by the increased cost of producing a diploid cell. In nutrient-limited environments this effect may be significant. In our experiments the average cell density for diploids was 3.4×10^7 cells ml⁻¹ during the early generations, whereas for haploids in the same environment the average cell density was 4.7×10^7 cells ml⁻¹. Thus the haploid population was larger by a factor of 1.4 and this increase in size may largely compensate for the reduced variation in fitness produced by haploids. In this respect it is interesting to note that marine yeasts, most of which are haploid, also live in a nutrient-limited environment⁵⁰. On the other hand it is commonly agreed that in laboratory cultures of *S. cerevisiae* on rich media containing both haploid and diploid cells, the diploid cells will eventually outgrow the haploid cells^{48,51} despite the fact that no physiological growth rate advantage for diploid cells has been detected⁶. Thus one factor which may determine

the evolution of diploidy is the type of environment in which the organism lives.

We thank J. Pringle and the Yeast Genetics Stock Center, Berkeley, California for providing strains, and R. Bender,

Received 27 August 1982; accepted 19 January 1983.

- Williams, G. C. *Sex and Evolution* (Princeton University Press, 1975).
- Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
- Svedelius, N. *Proc. int. Congr. Pl. Sci. Ithaca, N.Y.* 1, 457-471 (1929).
- Stebbins, G. L. *Variation and Evolution in Plants* (Columbia University Press, 1950).
- Crow, J. F. & Kimura, M. *Am. Nat.* 99, 439-450 (1965).
- Adams, J. & Hansche, P. E. *Genetics* 76, 327-338 (1974).
- Lewontin, R. C. *The Genetic Basis of Evolutionary Change* (Columbia University Press, 1974).
- Crow, J. F. & Kimura, M. *An Introduction to Population Genetics Theory* (Harper & Row, New York, 1970).
- Weiss, R. L., Kukora, J. R. & Adams, J. *Proc. natn. Acad. Sci. U.S.A.* 72, 794-798 (1975).
- Muller, H. J. *Am. Nat.* 8, 118-138 (1932).
- Sonneborn, T. M. in *The Biology of Aging* (eds Behnke, J. A., Finch, C. E. & Moment, G. B.) 361-374 (Plenum, New York, 1978).
- Bernstein, H., Beyers, G. & Michod, R. *Am. Nat.* 117, 537-549 (1981).
- King, J. L. & Jukes, T. H. *Science* 164, 788-798 (1969).
- Kimura, M. & Ohta, T. *J. molec. Evol.* 1, 1-17 (1971).
- Drake, J. W. in *Evolution in the Microbial World* (eds Carlile, M. J. & Skehel, J. J.) 41-58 (Cambridge University Press, 1974).
- Huxley, J. *Evolution: The Modern Synthesis* (Allen & Unwin, London, 1942).
- Dayhoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5 (National Biomedical Research Council, Silver Spring, Maryland, 1972).
- Jensen, R. A. *Rev. Microbiol.* 30, 409-425 (1976).
- McIntyre, R. J. A. *Rev. Ecol. Syst.* 7, 421-468 (1976).
- Hood, L., Campbell, J. H. & Elgin, S. C. R. *A. Rev. Genet.* 9, 305-323 (1975).
- Ohno, S. *Evolution by Gene Duplication* (Springer, Berlin, 1970).
- Lindsley, D. L. & Grell, E. H. *Carnegie Inst. Wash. Publ.* 627 (Washington D.C., 1968).
- McKusick, V. A. & Chase, G. A. *A. Rev. Genet.* 7, 435-473 (1973).
- Koch, A. L. *Genetics* 77, 127-142 (1974).
- Atwood, K., Schneider, L. & Ryan, F. J. *Cold Spring Harb. Symp. quant. Biol.* 16, 345-355 (1951).
- Maynard Smith, J. & Haigh, J. *Genet. Res.* 23, 23-35 (1974).
- Kubitschek, H. R. in *Evolution in the Microbial World* (eds Carlile, M. J. & Skehel, J. J.) 105-130 (Cambridge University Press, 1974).
- Paquin, C. E. & Adams, J. *Curr. Genet.* 6, 21-24 (1982).
- Lindgren, B. W. *Statistical Theory* (Macmillan, New York, 1976).
- W. D. Hamilton, J. Pringle and P. Smouse for helpful comments on an earlier version of this manuscript. This work was supported in part by NSF grant DEB 80-22185 and NIH grant S-T32-GM07544.
- Novick, A. & Szilard, L. *Proc. natn. Acad. Sci. U.S.A.* 36, 708-719 (1950).
- Novick, A. in *Perspectives in Marine Biology* (ed. Buzzati-Traverso, A. A.) 533-547 (University of California Press, 1960).
- Thorne, R. S. W. *J. Inst. Brew.* 74, 516-524 (1968).
- Thorne, R. S. W. *J. Inst. Brew.* 76, 555-563 (1970).
- Valenzuela, P., Medina, A., Rutter, W. J., Ammerer, G. & Hall, B. D. *Nature* 298, 347-350 (1982).
- Timmis, K. N. *et al. Molec. gen. Genet.* 167, 11-19 (1978).
- Adams, J., Kinney, T., Thompson, S., Rubin, L. & Helling, R. B. *Genetics* 91, 627-637 (1979).
- Isaacs, J. in *Perspectives in Marine Biology* (ed. Buzzati-Traverso, A. A.) 545 (University of California Press, 1960).
- Paquin, C. thesis, Univ. Michigan (1982).
- Holland, J. *et al. Science* 215, 1577-1585 (1982).
- Gibson, T. C., Scheppe, M. L. & Cox, E. C. *Science* 169, 686-688 (1970).
- Nestmann, E. R. & Hill, R. F. *Genet. Suppl.* 73, 41-44 (1973).
- Nestmann, E. R. & Hill, R. F. *J. Bact.* 119, 33-35 (1974).
- Painter, P. R. *Genetics* 79, 649-660 (1975).
- Kimura, M. *Genet. Res.* 9, 25-34 (1967).
- Eshel, I. *Theor. Populat. Biol.* 4, 196-208 (1973).
- Leigh, E. *Genet. Suppl.* 73, 1-18 (1973).
- Darlington, C. D. *The Evolution of Genetic Systems* 2nd edn (Basic Books, New York, 1958).
- Fowell, R. R. in *The Yeasts* Vol. 1 (eds Rose, A. H. & Harrison, J. S.) 461-471 (Academic, New York, 1969).
- Lewis, J. & Wolpert, L. *J. theor. Biol.* 78, 425-438 (1979).
- Van Uden, N. & Fell, J. W. *Adv. Microbiol. Sea* 1, 167-202 (1968).
- Roman, H., Phillips, M. M. & Sands, S. M. *Genetics* 40, 546-561 (1955).
- Helling, R. B., Kinney, T. & Adams, J. *J. gen. Microbiol.* 123, 129-141 (1981).
- Magni, G. & Von Borstel, R. *Genetics* 47, 1097-1108 (1962).
- Lea, D. & Coulson, A. J. *Genet.* 49, 264-285 (1949).
- Wilkie, D. & Lee, B. K. *Genet. Res.* 6, 130-138 (1965).
- Francis, J. C. & Hansche, P. E. *Genetics* 70, 59-73 (1972).
- Broach, J. R. in *The Molecular Biology of the Yeast Saccharomyces I*, 653-727 (Cold Spring Harbor Laboratory, New York, 1982).
- Kubitschek, H. E. *Introduction to Research with Continuous Cultures* (Prentice-Hall, New Jersey, 1970).
- Gottlieb, D. & Von Borstel, R. *Genetics* 83, 655-666 (1982).

Transmission of conformational change in insulin

Cyrus Chothia^{*†}, Arthur M. Lesk^{*||}, Guy G. Dodson[‡] & Dorothy C. Hodgkin[§]

^{*} Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[†] William Ramsay, Ralph Foster and Christopher Ingold Laboratories, University College London, Gower Street, London WC1E 6BT, UK

[‡] Department of Chemistry, University of York, Heslington, York YO1 5DD, UK

[§] Chemical Crystallography Laboratory, Parks Road, Oxford OX1 3PD, UK

Crystal structures of insulin contain molecules that are similar but not identical in conformation. Packed helices move relative to each other, these shifts being accommodated by motions of side-chain atoms arising from small changes in torsion angles. Such low-energy conformational adjustments can accommodate shifts of no more than ~1.5 Å. This limits the extent to which conformational changes can be dissipated locally, causing their transmission over long distances.

MANY properties of proteins, including the regulation of their activity, depend on transmission of conformational changes between distant sites. Little is known about the detailed mechanisms of these processes, as few protein structures are known to sufficiently high accuracy in more than one conformational state. Highly refined structures of several forms of insulin¹⁻³ reveal a similar fold of the chains but extensive differences in conformational details. Here we extend the previous comparisons of these structures⁴⁻⁶ to elucidate the mechanism of transmission of conformational changes in insulin.

We find that the backbones of helices tend to move as rigid bodies during conformational changes. Small relative displacements of packed helices (up to ~1.5 Å) are readily accommodated by small adjustments in torsional angles (up to ~15°) without the repacking of interfaces. This may be thought of as the limit of plastic deformation of helix interfaces. Larger displacements of packed helices are less easily accommodated, and occur very rarely. We suggest that it is the limited extent

to which internal motions can be dissipated locally that is responsible for the transmission of conformational change over long distances.

Although individual proteins have individual mechanisms of function, these ideas should apply generally to conformational changes in proteins, including allosteric transitions.

Structure of pig insulin

The insulin monomer contains two chains, A and B, containing 21 and 30 residues, respectively, and linked by two disulphide bridges, A7-B7 and A20-B19. A third disulphide bridge lies within the A chain at A6-A11. Two monomers form a dimer, held together by hydrogen bonding between two strands of antiparallel β -sheet, and by van der Waals' contacts.

In the presence of Zn²⁺ the dimers assemble into hexamers. Their conformation is solvent-dependent^{7,8}—crystallization at low ionic strength yields the 2Zn form, with two Zn²⁺ ions bound per hexamer; crystallization at high ionic strength yields the 4Zn form, which can bind four Zn²⁺ ions per hexamer. Both crystal forms are rhombohedral, with the crystallographic 3-fold axis relating the three dimers in the hexamer (Fig. 1a). The asymmetric unit of each crystal contains two monomers

^{||} Permanent address: Fairleigh Dickinson University, Teaneck-Hackensack Campus, Teaneck, New Jersey 07666, USA.

which are similar but not identical in conformation (Fig. 1). In the 2Zn form, two Zn^{2+} ions lie on the 3-fold axis, coordinated by His residues (Fig. 1b). In the 4Zn form, one Zn^{2+} occupies a site on the 3-fold axis, as in the 2Zn form; three others bind in off-axial positions between two dimers (Fig. 1c).

The crystal structures of 2Zn and 4Zn insulin have been determined to a resolution of 1.5 Å and refined to give crystallographic agreement factors (*R*) of 17.9% and 19.0%, respectively. The errors in position of nearly all the main-chain atoms and the buried side-chain atoms are less than 0.2 Å.

Figure 1 shows the secondary structures and disulphide bridges of 2Zn and 4Zn insulin. Each A chain contains two helices (A_I , residues A1–A9; and A_{II} , residues A12–A19). The B chains contain a single helix (residues B9–B19, except for one monomer in 4Zn insulin in which this helix extends to B1) and a strand of β -sheet (B24–B26) hydrogen bonded to the corresponding strand of the other monomer.

One monomer of the 2Zn form is very similar in conformation to the corresponding monomer of the 4Zn form and to both molecules in hagfish insulin⁶. We take this monomer as the 'reference structure' in our analysis of the conformational changes. Adopting the Chinese group's convention⁹, we call this monomer molecule 1, and the other monomer molecule 2.

Structural differences between insulin molecules

The fit of the two molecules in 2Zn insulin: We surveyed the two monomers of 2Zn insulin for segments that persist as nearly-rigid units in the two molecules. Certain regions of the backbone have the same conformation in both monomers, to within experimental accuracy. For these segments, the r.m.s. differences in atomic positions after superposition of the main-chain atoms (N–C α –C–O) are in the range 0.11–0.19 Å. Eight such segments were identified (Fig. 2a). They cover 90% of the A and B chains.

These are the largest rigid segments of the backbone, for if two or more of them are simultaneously superposed, the r.m.s. differences in atomic positions are substantially increased. This implies that despite the similar folding topology, the segments have moved relative to one another.

To describe quantitatively the shifts and rotations of the rigid segments, we overlaid molecules 1 and 2 by superposing the main-chain atoms of the B-helix segments (residues B9–B19) and then calculated the additional translation and rotation required to superpose other segments of the A and B chains. Figure 2b gives the changes in position and orientation of each structural segment, relative to the B helix. All segments show significant displacements.

The near-congruence of structure in segments of the main chains does not extend to the side chains. For residues on the surface this might be attributable to sensitivity to solvent environment. However, this explanation could not apply to the 15 well resolved side chains buried in the monomers (accessible surface area ≤ 20 Å²):

A2	A3	A6	A11	A12	A16	A20
Ile	Val	Cys	Cys	Ser	Leu	Cys
B6	B11	B12	B14	B15	B19	B24
Leu	Leu	Val	Ala	Leu	Cys	Phe
						Tyr

All these residues lie in segments for which the main-chain atoms of molecules 1 and 2 of 2Zn insulin fit with r.m.s. atomic distances of 0.11–0.19 Å.

After superposing the main chains, the r.m.s. difference in atomic position of the atoms of all but two of these buried side chains is 0.38 Å. The small differences in position of these side-chain atoms are the cumulative effect of small torsion angle changes of $\leq 13^\circ$.

Two buried side chains have very different conformations: internal rotations of $\sim 120^\circ$ move C γ 1 of A2 Ile and the C γ s of

Fig. 1 a, The hexameric packing of molecules in 2Zn and 4Zn insulin. The 3-fold axis that relates the dimers is perpendicular to the page. b, c, Schematic drawings of the 2Zn and 4Zn dimers. Helices are shown as cylinders and labelled A_I , A_{II} and B. β -sheet strands are shown as ribbons. Shaded circles represent sulphur atoms in disulphide bridges (the A6–B7 bridges are hidden by the A_I helices). In this part of the figure the 3-fold axis is vertical. The two monomers in the asymmetric unit of each crystal are related by an approximate 2-fold axis perpendicular to the page. In the 2Zn structure both zinc ions are necessarily located on the 3-fold axis. The 4Zn structure has one zinc ion on the 3-fold axis and the other three in off-axial sites between the three pairs of molecules 2. The figure shows molecule 2 in the two off-axial zinc ions that bind to 4Zn insulin.

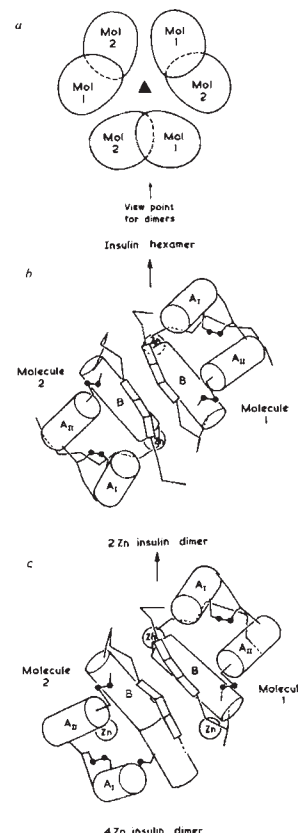


Table 1 Surface side chains involved in the movements of segments of the 2Zn insulin molecule

	B5 His		A19 Tyr		B25 Phe	
	χ_1	χ_2	χ_1	χ_2	χ_1	χ_2
Molecule 1	-82°	-76°	-51°	-89°	55°	-96°
Molecule 2	-179°	-137°	-68°	-89°	-91°	-97°

B12 Val by ~ 2 Å. Other, less buried, side chains also show large conformational differences. Table 1 shows side chains of importance in understanding the relationships among the movements of different segments (Fig. 4a).

In summary, the two monomers of 2Zn insulin can be dissected almost entirely into segments of very similar main-chain conformation, which differ in their relative spatial disposition. The side chains which form the contacts between these segments differ in conformation, usually to only a small extent.

The fit of 4Zn molecule 2 to 2Zn molecule 1: The conformations of molecule 1 and the axial Zn binding site are very similar in 2Zn and 4Zn insulin (Fig. 1). The major structural difference between 2Zn and 4Zn insulin is the refolding of the N-terminus of the B strand of molecule 2. In other insulin monomers, residues B1–B8 form a nearly-extended chain. In molecule 2 of 4Zn insulin, residues B1–B8 shift by up to 21 Å to form an α -helix contiguous with the helix formed in other monomers by residues B9–B20 (Figs 1, 3). This allows the two His side chains to take up the proper relative positions to form new Zn coordination sites. The three off-axial Zn^{2+} ions bind between dimers, to B5 His of each molecule 2 and B10 His of an adjacent molecule 2 (Fig. 1).

Five segments of the backbone of 4Zn molecule 2 have conformations similar to the corresponding parts of the reference molecule, 2Zn molecule 1 (r.m.s. differences in atomic position 0.16–0.25 Å (Fig. 2a)). These well fitting segments contain 30 residues out of a total of 51 in the monomer,

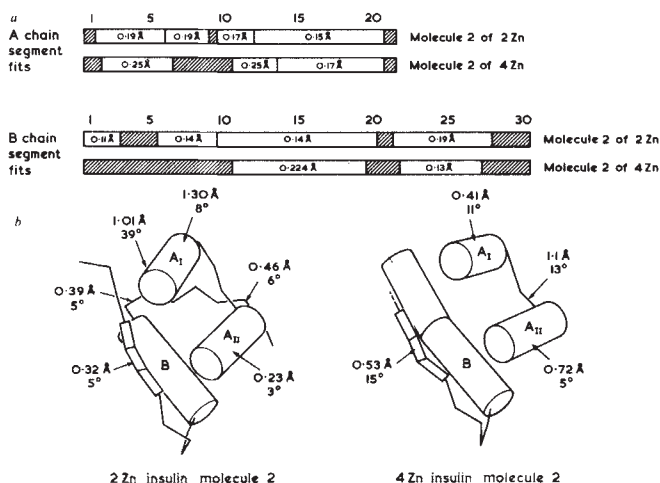


Fig. 2 The fits and relative shifts of main-chain segments in 2Zn and 4Zn insulins. *a*, The strips represent the A and B chains of molecules 2 of 2Zn and 4Zn insulin. Segments whose coordinates can be superposed well onto the homologous regions of molecule 1 of 2Zn insulin are shown as open regions, with the r.m.s. difference in atomic position reported. Regions that do not fit well are shaded. *b*, Molecules 2 of 2Zn and 4Zn insulin can be superposed on the reference structure, molecule 1 of 2Zn insulin, by fitting the main-chain atoms of the B helix (see text and Fig. 4). Because of the differences in conformation, other segments have shifted and rotated with respect to the homologous regions of the reference structure. We indicate in the figure the magnitudes of the translation and rotation necessary to superpose homologous segments.

including all the regions of common secondary structure except for the C-terminus of the A₁ helix (A7–A9).

As in 2Zn insulin, the buried side chains show significant differences in conformation. For most of the buried side chains, torsion angles change by no more than 14°, producing a r.m.s. difference in atomic positions of 0.43 Å. However, large changes are observed for B11 Leu and for the cysteines A7, A11 and B7. For these residues χ values change by 100°–170° and the side-chain atoms shift by 1–2 Å. These changes are required by the refolding of B1–B8. On the surface of the molecule, χ_1 of B25 Phe changes to –66°, giving this side chain a conformation similar to that in molecule 2 of 2Zn insulin (Fig. 4).

The common structural segments of molecule 2 of 4Zn insulin and molecule 1 of 2Zn insulin have different relative spatial arrangements. Figure 2*b* shows the shifts and rotations, after superposition of the B1 helices of molecule 2 of 4Zn and molecule 1 of 2Zn. The shifts are between 0.4 and 1.1 Å.

The monomers pack together differently in 2Zn and 4Zn insulin. We superposed the B helices and β -strands of molecules 1 of 2Zn and 4Zn insulin. These form one side of the dimer interface. On the other side, the B helices of molecules 2 differ in position by 1.1 Å and in orientation by 10.1° (Fig. 3). The β -strand of molecule 2 of 4Zn insulin, restrained by the β -sheet hydrogen bonding to molecule 1, does not follow the shift of the B helix.

Relationships between the structural differences

Structural differences extend throughout the insulin molecules. In this section we discuss the relationships among the conformational changes, and suggest that they can be understood, at least in part, as the propagation of the response to an initial perturbation localized at the molecular surface. We prefer this view to the alternative, that the changes arise independently, for two reasons. First, we can identify a single perturbation at the surface of each molecule that is the only significant difference in the molecular environment, and which requires that some conformational change take place. Second, our

hypothesis leads to a model for the flow of conformational change through the molecule that is logical and plausible. The initial perturbation requires adjustments in the structures of neighbouring segments, which in turn induce changes in the next neighbours.

Our descriptions of the propagation of conformational changes, given below, are based on an interpretative study of observed structures. Eventually, computational techniques such as molecular dynamics should be able to test our conclusions¹⁰.

Two-zinc insulin

Structural differences between molecules 1 and 2 of 2Zn insulin originate at the intermolecular contact between hexamers involving the end of the A₁ helix and residue B5 His of molecule 1 (Figs 4*a*, 5).

(1) **The geometrical requirements of the intermolecular contacts formed by the A₁ helix require this helix to shift.** The side chain of B5 His of molecule 1 forms hydrogen bonds with both the A₁ helix of its own monomer and the A₁ helix of molecule 2 of a neighbouring hexamer (Fig. 5*a*, left). These hydrogen bonds, and van der Waals' contacts between the A10 Ile side chains, constrain the relative positions of the two A₁ helices: the A₁ helix of molecule 2 is shifted by 1.3 Å from its position in molecule 1 (Fig. 2*b*).

(2) **Impediments to the shift of the A₁ helix produce a conformational change at residue A6.** Inter-hexamer contacts shift the C-terminal residues of helix A₁ of molecule 2. The entire A₁ helix cannot follow this shift as a rigid body, because a collision would occur between A2 Ile and B11 Leu (Fig. 5*b*). The N-terminal segment of this helix rotates by 33° about ϕ of A6, changing this part of the helix from π to α conformation. The shift and rotation of the A₁ helix breaks the salt bridge A1 NH₃⁺...O₂C B30 which occurs in 2Zn molecule 1.

(3) **The shift of the A₁ helix is accommodated by changes in side-chain conformation in residues of helix A₁₁, which packs against it.** On the A₁₁ side the biggest change is a χ_1 rotation of 17° in A19 Tyr, which displaces its hydroxyl oxygen by 1.5 Å (Figs 4, 5*b*).

(4) **The movement of A19 Tyr ejects B25 Phe from its pocket inside the molecule.** In molecule 1 the ring of B25 Phe packs into its own monomer, between A19 Tyr and B27 Thr (Fig. 4*a*). This is impossible in molecule 2 because A19 Tyr has taken the space that B25 Phe would occupy (Fig. 5*c*). In molecule 2 the ring points across the dimer–dimer interface to pack against B25 Phe of the other monomer.

Note that B25 Phe is 20 Å from the site of the initial perturbation at the end of the A₁ helix.

Four-zinc insulin

The transition from 2Zn to 4Zn insulin is produced by a change in ionic strength of the solution rather than by the specific effects of Zn²⁺ binding, and indeed the change in backbone conformation can occur without the uptake of additional zinc⁶. (The threshold salt concentration for the transition depends on the anion according to the Hofmeister series⁸, suggesting that it is a change in the relative magnitude of hydrophobic and nonhydrophobic interactions which is responsible.) We shall begin our analysis of the conformational changes with the refolding of the N-terminus of the B chain of molecule 2, and describe its effects.

The new interactions of the N-terminal extension of the B helix cause it to shift with respect to the dimer interface. The refolded segment B1–B8 has several interactions unique to molecule 2 of 4Zn insulin: (1) It packs against the extended B helices of the other two molecules 2 in the hexamer. (2) It packs against the carbonyl end of the A₁ helix of molecule 2 in an adjacent dimer. (3) Within its own molecule it alters the B chain–A₁ helix interactions.

The major effects of these new B-helix contacts are to change the dimer interface and to shift the A chain with respect to the B within 4Zn molecule 2.

At the dimer interface, the C-terminus of the B helix of molecule 2 (B9–B19, the residues common to B helices of other monomers) packs against the B helix of molecule 1 and the two β -strands. Comparison of the dimer interfaces of 2Zn and 4Zn insulin reveals a 1.1 Å shift in the B helix of molecule 2 (Fig. 3). The shift is accommodated by torsion angle changes of up to 15° in residues forming the dimer interface.

The extension of the B helix causes the A_I helix to move, both through direct steric interaction and through changes in hexamer–hexamer contacts. The extension of the B helix is in contact with the C-terminus of the A_I helix of molecule 2 of an adjacent dimer in the same hexamer (Fig. 5a). The presence of the B-helix extension at the surface of the hexamer also requires changes in hexamer–hexamer contacts involving the A_I helix. In 2Zn insulin, there is an inter-hexamer hydrogen bond between B5 His of molecule 2 and the carbonyl group of A9, and van der Waals' contact between A10 Ile residues of adjacent hexamers (Fig. 5a)⁵. In 4Zn insulin, interaction of the three extended B helices prisms the hexamers 3.6 Å further apart. The B5 His hydrogen bonds to A8 rather than A9 and the A chains are not in contact.

The formation of the extended helix also has the effect of removing the contacts that occur in molecule 1 between B1–B5 and this part of the A chain.

As a result, residues A8–A10 shift by ~3.0 Å and change their conformation (Figs 4, 5a). (These residues include the C-terminus of the A_I helix and part of the turn between helices A_I and A_{II}.) The large shift is partially dissipated by local changes in backbone conformational angles, but some of it is transmitted to the A helices. The N-terminal part of the A_I helix, and the A_{II} helix, move without significant deformation of their backbones away from the B helix and β -strand (Fig. 4b). Helix A_{II} shifts, relative to the B helix, by 0.72 Å, and rotates by 4.9°. The packing accommodates these motions by torsion angle changes of $\leq 15^\circ$.

The shift of the A_{II} helix loosens the packing of the side chain of B25 Phe, which rotates to point out across the dimer interface. The shift of the A_{II} helix away from the B strand produces a conformational change in residue B25 Phe. In 2Zn molecule 1, this residue packs into its own monomer, making contact with residues A19 Tyr and A21 Asn. The shift in the A helices in 4Zn molecule 2 moves A19 Tyr approximately 1 Å away from the B strand. If the side chain of B25 Phe stayed in the same position it would sit within a cavity, packing relatively poorly. Instead, it rotates to point out across the monomer–monomer interface (Fig. 4b).

It is interesting that, for opposite reasons, the same conformational change of this ring occurs in both 2Zn and 4Zn insulin (Fig. 5c): in 2Zn B25 Phe is squeezed out of a pocket that is too tight; in 4Zn it drops out of a pocket that is too loose.

General features of the conformational changes in insulin

Although the ends of chains and the regions between secondary structures are relatively flexible and can, on suitable provocation, entirely refold, well formed α -helices shift by up to ~1.3 Å with almost no significant change in backbone conformation. For the four monomers in the 2Zn and 4Zn crystals, the first part of the A_I helix, the A_{II} and B helices and the β -strands can be superposed with r.m.s. differences in atomic positions of 0.11–0.24 Å. Conformational changes within the helices, when they do occur, tend to be localized rather than spread out; that is, helices 'kink' rather than bend smoothly.

The nearly rigid displacements of packed helices are facilitated by side-chain atom motions in the interfaces. Small torsion angle changes, usually less than 10° and rarely more than 15°, can have the cumulative effect of shifting side-chain atoms by up to ~1.0 Å. Very occasionally, large torsion angle changes shift buried side-chain atoms by 1–2 Å.

With one exception, the relative movements of the main chains of the packed helices are no greater than 1.3 Å. These

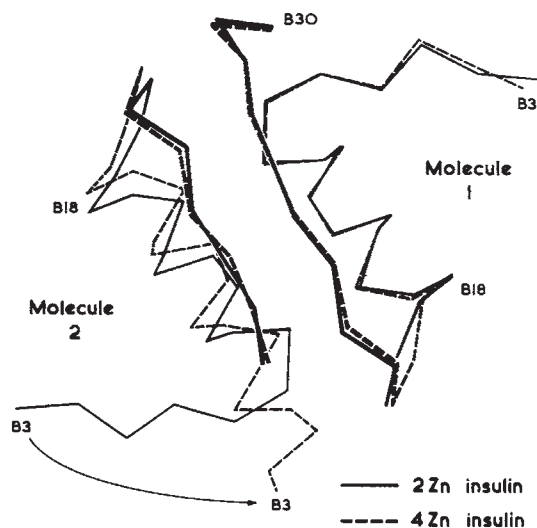


Fig. 3 Differences in 2Zn (solid line) and 4Zn (broken line) insulin at the monomer–monomer interface. The interface is formed by the B helices and the β -strands. The 2Zn and 4Zn dimers are superposed by fitting the B helices and β -strands of molecule 1. The B helices of molecule 2 differ in position by 1.1 Å and in orientation by 10°.

shifts do not require repacking of interface residues: the pattern of residue–residue contacts is conserved. However, the ~2 Å shift of the A_I helix relative to A_{II} in molecule 2 of 2Zn insulin does require repacking of side chains (Fig. 3b).

Helix movements in proteins

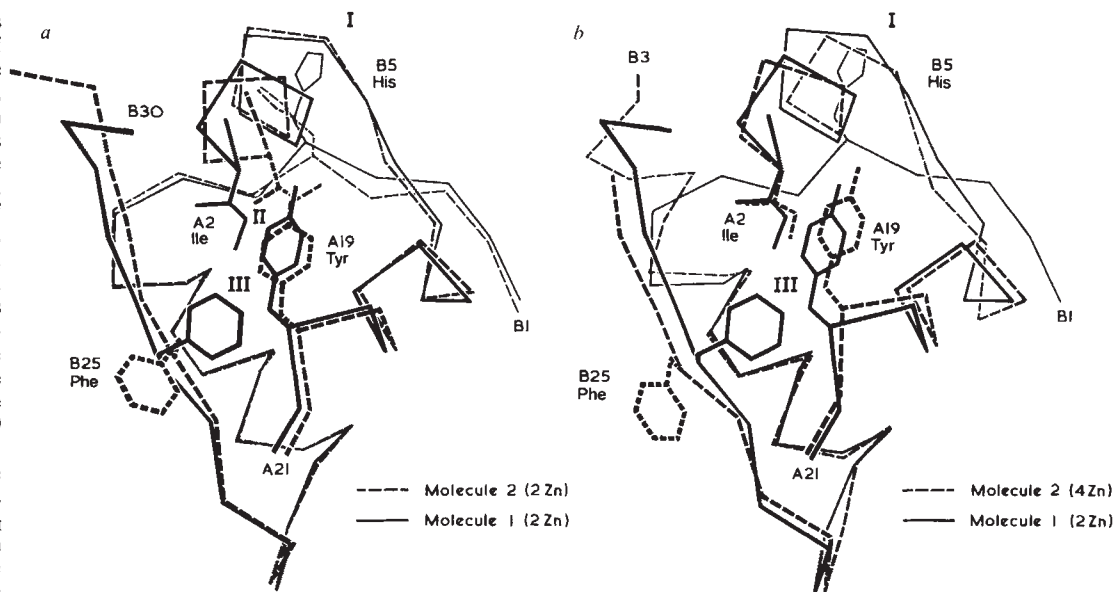
Because shifts in packed helices tend to be relatively small (~1 Å), accurately determined structures such as insulin are necessary for a detailed investigation of the mechanism of transmission of conformational changes. Studies of other proteins have described the much larger changes in folding of loops and chain termini, which are important for the function of haemoglobin¹¹ and lactic dehydrogenase¹², and for the activation of trypsin¹³. Detailed information has not previously been available for helix movements. In adenylate kinase¹⁴ two adjacent helices shift by ~1–2 Å. In haemoglobins¹⁵ the allosteric transition involves helix shifts of up to 1.6 Å. For both these proteins, the resolution of the structures being compared, 3.0 Å and 4.7 Å in the first case and 2.0 Å and 2.8 Å in the second, prevented an accurate description of how the shifts were accommodated. Using NMR, Williams and co-workers have described helix movements produced by changes of ligation in calcium-binding proteins¹⁶.

The insulin structures have shown clearly the nature of the structural change involved in movements of packed helices. The main chain remains comparatively rigid; the side chains undergo small internal rotations. The comparative rigidity of main chains of the α -helices must arise from the geometrical requirements of their hydrogen bonds and the narrow steric constraints on the backbone conformational angles.

The insulin structures do not provide information about the transmission of conformational change by β -sheets. (Because the allowed range of (ϕ , ψ) values is larger for β -sheets than for α -helices, and the β -sheet structure is intrinsically less compact, β -sheets might more readily dissipate conformational changes locally.)

Side-chain internal rotations accommodate and transmit the relative displacements of packed helices. The residues which transmit conformational changes between helices make firm contacts with their neighbours, either through the hydrogen bonding of the helix, or by being well packed in the protein interior. These atoms have low temperature factors (B), and in general are not adjacent to any cavity in the molecule. Therefore, accommodation of displacements of packed helices requires the adjustment of side-chain conformations. The limit

Fig. 4 Relationships between the shifts of the segments in the insulin structures. In *a*, molecule 2 of 2Zn insulin (broken line) is superposed on molecule 1 of 2Zn insulin (solid line). In *b*, molecule 2 of 4Zn insulin (broken line) is superposed on molecule 1 of 2Zn insulin (solid line). This figure shows line segments joining successive C α s in the polypeptide backbones, and the side chains of A2 Ile, A19 Tyr, B5 His and B25 Phe. *a*, In region I the A₁ helix of molecule 2 interacts with the A₁ helix of molecule 1 of an adjacent hexamer (see Fig. 5a, opposite). The B5 His side chain makes



hydrogen bonds to residues of both helices, and there are van der Waals' contacts between residues A10 Ile. These crystal-packing interactions require that the A₁ helices of molecules 1 and 2 adopt different positions: the molecule 2 helix is shifted by ~1 Å in a direction into the page. In region II the A₁ helix movement affects the A₁-A₁₁ helix packing. The N-terminal part of the A₁ helix has shifted by 1 Å and rotated by 37°. This changes the packing of A2 Ile against A19 Tyr (see Fig. 5b, opposite). The shift of the A₁ helix also breaks the salt bridge between A1 and B30. The change in packing of A19 Tyr moves it into part of the space occupied in molecule 1 by B25 Phe. Region III shows the resulting conformational change of B25 Phe (see Fig. 5c). *b*, In molecule 2 of 4Zn insulin residues B1-B8 refold to create an extension of the helix formed by residues B9-B19 (region I). The helix extension B1-B8 occupies space near the 3-fold axis. It packs against the corresponding portions of the other molecules 2 in the hexamer, and also against the end of the A₁ helix of an adjacent monomer (see Fig. 5a, opposite). These interactions cause a conformational change in residues A8-A9 (see Fig. 5a) and an overall shift of the helix. This shift moves A19 Tyr so as to loosen its contact with B25 Phe which rotates out of the monomer (see Fig. 5c). It is interesting that residues B25 Phe of both 2Zn and 4Zn molecules 2 show similar conformational changes, relative to 2Zn molecule 1, but for different reasons.

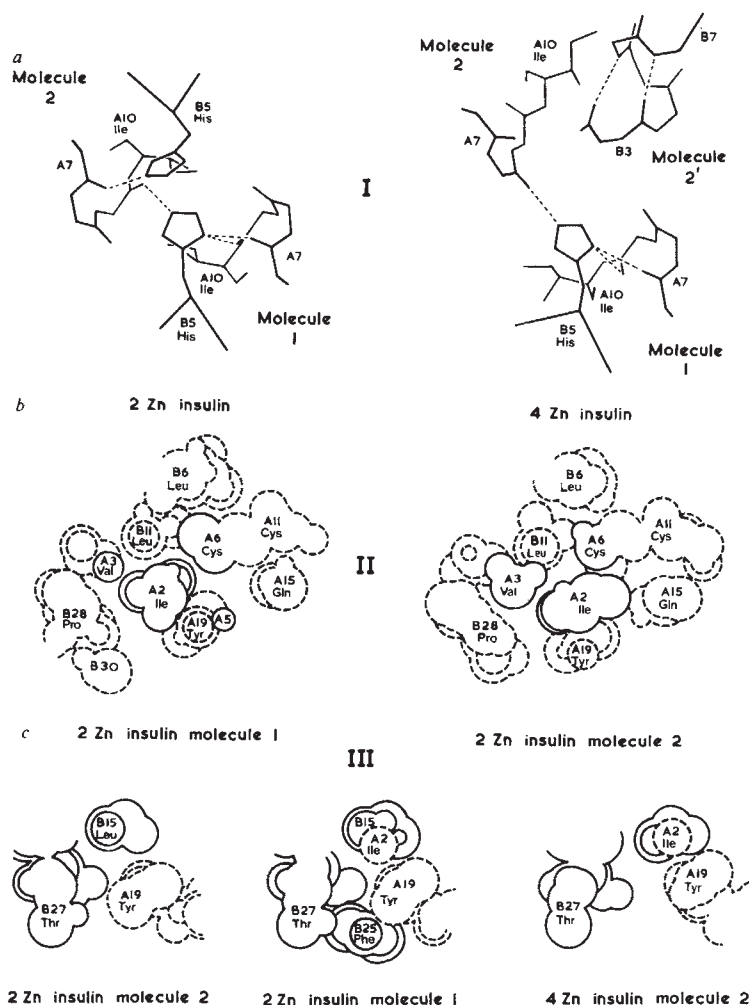


Fig. 5 Structural differences among insulin molecules. This figure shows the regions labelled I, II and III in Fig. 4 in greater detail. *a*, Interhexamer contacts at the end of the A₁ helix in 2Zn and 4Zn insulin. For 2Zn insulin (left) the figure shows main-chain atoms of residues A7-A10 of helix A₁ of molecule 1, the corresponding atoms of helix A₁ of molecule 2 of an adjacent hexamer, B5 His of molecule 1 which forms hydrogen bonds to both A₁ helices, and B5 His of molecule 2 which forms a hydrogen bond to the carbonyl group of A7. For 4Zn insulin (right) the two A₁ helices and the B5 His of molecule 1 are shown, together with the main-chain atoms of residues B3-B7 of an adjacent molecule 2 (molecule 2'), which have refolded into an extension of the B helix. (Residue B1 of molecule 2 and part of B2 are disordered in 4Zn insulin.) This refolding removes B5 His of molecule 2 from its position shown in 2Zn insulin (left). The two forms of insulin differ in the hydrogen-bonding pattern by which B5 His of molecule 1 interacts with the two A₁ helices. In 2Zn insulin this side chain forms hydrogen bonds with the carbonyl oxygens of A7 and A9 of its own monomer and with that of A9 of molecule 2 in an adjacent hexamer. There are also van der Waals' contacts between the A10 Ile of the two molecules 2. This packing arrangement cannot occur in the 4Zn insulin structure because the extended B helix in one molecule 2 packs against the end of the A₁ helix in an adjacent molecule 2. In 4Zn insulin the hexamers are 3.6 Å further apart and the inter-hexamer hydrogen bond is formed from B5 His to the carbonyl oxygen of A8 instead of A9. *b*, The packing of the A₁ helix in molecules 1 and 2 of 2Zn insulin. The first part of the A₁ helix, residues A1-A6, differs in position in the two molecules by ~2 Å (see text). This difference involves repacking by A3 Val, A2 Ile and A19 Tyr. *c*, The B25 Phe surface cavity. In molecule 1 of 2Zn insulin (centre), the ring of B25 Phe packs into a cavity in its own monomer between B27 Thr and A19 Tyr which are 8.6 Å apart. In molecules 2 the Phe side chain points out across the dimer interface (see Fig. 4). In molecule 2 of 2Zn insulin (left) the size of the cavity is too small: the distance between A19 and B27 is 6.3 Å. In molecule 2 of 4Zn insulin (right) the cavity is too large: the distance between A19 and B27 is 9.3 Å. (In molecule 1 it is just right.)

of allowable changes in side-chain conformation in turn limits the excursion of a packed helix that can occur without repacking the interface.

Our observation of limits to side-chain movements of buried residues is consistent with theoretical descriptions of their potential surfaces. Gelin and Karplus¹⁷, in their computational study of bovine pancreatic trypsin inhibitor, report that side-chain-main-chain interactions create a number of low-energy minima for the conformation of any residue. Interactions with the rest of the protein select one of these minima and sharpen it. Van der Waals' repulsions erect high barriers about the minima: a range in energy of 0.5 kcal mol⁻¹ corresponds to a range of torsion angle of 20° or less¹⁷.

The transmission of conformational changes

Packed α -helices can readily accommodate small shifts by small changes in torsion angles, together with an occasional large side-chain rotation to a different local minimum. (This is a phenomenon which evolution may use in designing proteins which for functional reasons must undergo conformational change and, conversely, represents a pitfall which proteins that

Received 6 December 1982; accepted 10 February 1983.

1. Blundell, T. L. *et al.* *Nature* **231**, 506–511 (1971).
2. Bentley, G. A., Dodson, E. J., Dodson, G. G., Hodgkin, D. C. & Mercola, D. A. *Nature* **261**, 166–168 (1976).
3. Cutfield, J. F. *et al.* *J. molec. Biol.* **132**, 85–100 (1979).
4. Dodson, E. J., Dodson, G. G., Hodgkin, D. C. & Reynolds, C. D. *Can. J. Biochem.* **57**, 469–479 (1979).
5. Dodson, E. J., Dodson, G. G. & Hodgkin, D. C. in *Frontiers of Bioinorganic Chemistry and Molecular Biology* (ed. Ananenko, A. N.) 145–150 (Pergamon, Oxford, 1980).
6. Cutfield, J. F. *et al.* in *Structural Studies on Molecules of Biological Interest* (eds Dodson, G., Glusker, J. P. & Sayre, D.) 527–546 (Clarendon, Oxford, 1981).
7. Schlichtkrull, J. in *Insulin Crystals* (Munksgaard, Copenhagen, 1958).

must remain rigid must be designed to avoid.) Larger shifts would require larger conformational changes of individual residues, or the repacking of interfaces. Repacking could occur rarely at interfaces of considerable size because the packed residues would have to form a completely new set of complementary interactions. This is analogous to the behaviour of a material that can deform smoothly up to a point beyond which it breaks.

This limit to the size of the shifts of packed α -helices limits the extent to which a conformational perturbation can be accommodated locally. It is the existence of a limit to the local dissipation of helix movements that can propagate over long distances—and in some cases can even amplify—conformational changes in protein structures.

We thank John Cresswell for the figures and our Cambridge and York colleagues for their comments on an earlier version of the paper. This work is supported by the Royal Society, NSF (PCM80-12007), the National Institute of General Medical Sciences (GM 25435), the North Atlantic Treaty Organization (RG275.80), EMBO (SF4-1980), the British Diabetic Association and the Wellcome Trust.

8. De Graaff, R. A. G., Lewit-Bentley, A. & Tolley, S. P. in *Structural Studies on Molecules of Biological Interest* (eds Dodson, G., Glusker, J. P. & Sayre, D.) 547–566 (Clarendon, Oxford, 1981).
9. The Beijing Insulin Structure Research Group in *Structural Studies on Molecules of Biological Interest* (eds Dodson, G., Glusker, J. P. & Sayre, D.) 501–508 (Clarendon, Oxford, 1981).
10. Delhaise, P. & Wodak, S. J. in *Non-bonded Interactions and the Specificity of Protein Association and Folding* (ed. Wodak, S. J.) 122–160 (CECAM, Orsay, 1981).
11. Fermi, G. & Perutz, M. F. *Haemoglobin & Myoglobin* (Clarendon, Oxford, 1981).
12. White, J. L. *et al.* *J. molec. Biol.* **102**, 759–779 (1976).
13. Huber, R. & Bode, W. *Acc. chem. Res.* **11**, 114–122 (1978).
14. Sachsenheimer, W. & Schulz, G. E. *J. molec. Biol.* **114**, 23–36 (1977).
15. Baldwin, J. & Chothia, C. *J. molec. Biol.* **129**, 175–220 (1979).
16. Levine, B. A. *et al.* *Ciba Fdn Symp.* **93**, 72–97 (1983).
17. Gelin, B. R. & Karplus, M. *Biochemistry* **18**, 1256–1268 (1979).

LETTERS TO NATURE

The impossibility of a bouncing universe

Alan H. Guth

Center for Theoretical Physics, Laboratory for Nuclear Science and Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Marc Sher

Department of Physics, University of California, Irvine, California 92717, USA

Petrosian¹ has recently discussed the possibility that the restoration of symmetry at grand unification in a closed contracting Robertson–Walker universe could slow down and halt the contraction, causing the universe to bounce. He then went on to discuss the possibility that our universe has undergone a series of such bounces. We disagree with this analysis. One of us (M.S.) has already shown² that if a contracting universe is dominated by radiation, then a bounce is impossible. We will show here two further results: (1) entropy considerations imply that the quantity S (defined in ref. 1 and below), which must decrease by $\sim 10^{75}$ to allow the present Universe to bounce, can in fact decrease by no more than a factor of ~ 2 ; (2) if the true vacuum state has zero energy density, then a universe which is contracting in its low temperature phase can never complete a phase transition soon enough to cause a bounce.

The scale factor $a(t)$ of a closed Robertson–Walker universe evolves according to³

$$\dot{a}^2 = Ea^2 - 1 \quad (1a)$$

$$\ddot{a} = -\frac{1}{2}(E + 3P)a \quad (1b)$$

where E is the energy density, P is the pressure, and the overdot

denotes differentiation with respect to time t . (We use the notation of ref. 1, with $8\pi G/3 = \hbar = c = 1$.) In ref. 1 the equations are written as

$$\begin{aligned} (a\dot{a})^2 &= E\dot{a}^4 - a^2 + S^{4/3} \\ \ddot{a}a^3 &= E\dot{a}^4 - S^{4/3} \end{aligned} \quad (2)$$

which agree with equation (1) provided that we adopt the definitions

$$E_V \equiv \frac{1}{4}(E - 3P) \quad (3a)$$

and

$$S^{4/3} \equiv \frac{3}{4}(E + P)a^4 \quad (3b)$$

From equation (2) it follows that

$$\dot{S}^{4/3} = -a^4 \dot{E}_V \quad (4)$$

In the scenario of ref. 1, the universe is contracting in its low-temperature phase, with $S = S_1$ ($\sim 10^{90}$ for the present Universe) and $E_V = E_1$ ($E_1 < 10^{-120}$ for the present Universe). The universe then undergoes a phase transition at $a = a_1$ associated, for example, with the restoration of the grand unified symmetry; we use the subscript t to denote the values of the various quantities after the phase transition. One then has $E_V = E_t \sim 10^{-20}$ and $S^{4/3} = S_t^{4/3} \approx S_1^{4/3} - a_1^4(E_t - E_1)$. Then taking E_V and S as constants, it is found that equations (3) predict a bounce if $4S_t^{4/3}E_t < 1$, or $S_t \lesssim 10^{15}$.

We agree with all the statements above, but we claim that entropy considerations preclude the possibility that S_t is small. To see this, let us first assume that thermal equilibrium is maintained throughout the transition. We will define $\bar{E}_V(T)$ to be the value of the zero-temperature Higgs potential, evaluated at the expectation value of the Higgs field at temperature T . ($\bar{E}_V(T)$ is sometimes called the 'vacuum energy density'.) Following ref. 1, we then write the pressure as

$$P = -\bar{E}_V(T) + \frac{\pi^2}{45}[1 + g(T)]T^4 \quad (5)$$

where $g(T)$ can be interpreted as the number of effectively massless degrees of freedom (other than photons) at temperature T . The quantities $\bar{E}_V$ and g have approximately constant values within each phase, but they will vary rapidly during the phase transition. Thermodynamics requires⁴ that if P is given as a function of T (we are considering a case in which all chemical potentials vanish), then

$$E = -P + T \frac{\partial P}{\partial T} \quad (6)$$

It can be seen that $E_V \approx \bar{E}_V$ whenever g and $\bar{E}_V$ are slowly varying with T .

The volume of the universe is $2\pi^2 a^3$, and so the total entropy⁴ is given by $\Sigma = 2\pi^2 a^3 \partial P / \partial T$. If $\bar{E}_V$ and g are both constant (which is a very good approximation except during the phase transition), then

$$S = 0.042(1+g)^{-1/4} \Sigma \quad (7)$$

We conclude that if thermal equilibrium is maintained during the phase transition, then Σ is conserved and S can decrease only due to a change in the factor $(1+g)^{-1/4}$; in a typical grand unified theory this factor changes by less than factor of 2. Thus, under these circumstances a bounce is precluded—on this point we agree with ref. 1. A bounce would be allowed only if $4S_1^{4/3} E_i \leq 2$.

Now let us consider the more plausible case in which the phase transition is delayed by superheating, resulting in a departure from thermal equilibrium. In this case entropy is not conserved, but the second law of thermodynamics implies that the entropy must increase. Thus, if equilibrium is restored in the new phase, then equation (7) will hold with a value of Σ which is greater than or equal to the value it had before the phase transition; we again conclude that S cannot decrease by more than a factor of about 2. (Note that equation (4) remains valid in this context; the second law of thermodynamics simply places a constraint on how rapidly E_V can change. In other words, the second law implies that the phase transition will not take place until the system has been compressed enough so that the entropy of the high-temperature phase is greater than or equal to that of the low-temperature phase.)

Finally, one might assume that thermal equilibrium is never restored in the new phase, invalidating the above argument. However, this assumption also invalidates the equation of state used in ref. 1 (that is, the assumption that E_V and S have constant values after the phase transition becomes unjustified). If thermal equilibrium is not restored, then even the concept of temperature becomes inapplicable. While we cannot rule out the possibility of a bounce under these circumstances, we believe that Petrosian¹ has not presented a consistent scenario which predicts such a bounce.

We now turn to our second and stronger result, which shows that even if $4S_1^{4/3} E_i < 1$, the phase transition cannot be completed soon enough to cause a bounce. Our argument will again allow for superheating and a departure from thermal equilibrium, but we will again assume that thermal equilibrium is restored in the high-temperature phase. No approximations for the equation of state will be necessary.

To explain this result, we first note that any region which is driven to the symmetric phase as the universe contracts will necessarily be at a temperature $T \geq T_0$, where T_0 is the critical temperature of the phase transition. (This is an analogous situation to what happens when a small amount of heat is added to a closed vessel containing water at the critical temperature T_0 of the water-steam phase transition. If the water turned entirely to steam, absorbing a large amount of latent heat, then conservation of energy would imply that the steam would be supercooled ($T < T_0$); however, this is not what happens. Instead bubbles of steam form; if the heating is slow, then T remains at T_0 , which increases as the pressure increases. If heat is added quickly one can obtain superheated water, but never supercooled steam.)

We next note that equation (1b) implies that the bounce can occur only if

$$E + 3P < 0 \quad (8)$$

This inequality can hold in the (high-temperature) symmetric phase ($E_V > 0$) at sufficiently low temperatures, but we will show that it cannot hold for $T > T_0$. We will assume for this argument that the true (zero-temperature) vacuum state of the theory has zero energy density. (This statement is certainly true to a high degree of accuracy.) It follows that the total energy of any other state is positive, and hence the energy density of any homogeneous state obeys $E \geq 0$. Now use the fact that for $T > T_0$, the symmetric phase is the thermal equilibrium phase. The pressure of the thermal equilibrium phase is given by⁴

$$e^{PV/T} = \text{Tr} \{ e^{-(H - \mu_i N_i)/T} \} \quad (9)$$

where Tr denotes the quantum mechanical trace, H is the hamiltonian of the system, and μ_i denotes the chemical potential for the conserved quantity N_i . (We no longer require the μ_i to vanish.) The vacuum contributes 1 to the right-hand side, and all terms are positive. We thus arrive at the very general result that $P \geq 0$ for any thermal equilibrium state. (As an alternative proof of this statement, note that any state with $P < 0$ is unstable to the formation of bubbles of true vacuum; since the true vacuum has higher pressure, any bubble of true vacuum which is large enough to overcome the effects of surface tension will expand without limit.)

Thus, we conclude that after the phase transition takes place (at $T \geq T_0$), inequality (8) will fail and it will be too late for the universe to bounce; it will instead continue to contract as long as the equations of classical general relativity remain valid.

We thank J. Goldstone for helpful comments. A.H.G. is supported in part through funds provided by the US Department of Energy under contract DE-AC02-76ERO3069 and an Alfred P. Sloan Fellowship. M.S. is supported in part through funds provided by the NSF under contract PHY-79-10262.

Received 4 October, accepted 30 December 1982.

1. Petrosian, V. *Nature* **298**, 805 (1982).
2. Sher, M. *Phys. Rev. D* **22**, 2989 (1980).
3. Weinberg, S. *Gravitation and Cosmology* (Wiley, New York, 1972).
4. Huang, K. *Statistical Mechanics* Chs 8, 9 (Wiley, New York, 1963).

Reply by Vahé Petrosian

Institute for Plasma Research, Stanford University, Stanford, California 94305, USA

First I shall comment on the paper by Guth and Sher¹ and then clarify the scenario for a bouncing universe described in my paper². After much discussion, correspondence and revision of previous writings and views, my ideas on this subject and those of Guth and Sher have converged to a great extent, but not completely. The remaining differences have to do with the degree of emphasis given to various scenarios and possibilities. One example is the title of the Guth-Sher paper, 'The impossibility of a bouncing universe', which contrasts with the statement in the text that the possibility of a bounce cannot be excluded if equilibrium is not maintained and re-established during the transition and bounce. I shall return to this below and present scenarios which show that conditions for a bounce can be satisfied and that such strongly worded conclusions are not warranted. Aside from this and other minor points (such as my belief that the water-steam analogy is not an appropriate one), I agree with the general conclusions of Guth and Sher, given their assumptions.

They have shown that if during a phase transition thermal equilibrium is maintained and, more importantly, if the equi-

brum is not maintained but is re-established on a time scale shorter than the gravitational time scale (the Hubble time at that epoch), a bounce becomes impossible. The second more general argument is identical to Sher's³ earlier argument with which I agreed². However, Guth and Sher's arguments are more general than that given by Sher as they include the effects of changes in the degrees of freedom of massless particles and of non-zero chemical potentials.

All these arguments rely on thermodynamic relations valid in thermal equilibrium and on the assumption that the temperature remains above the critical temperature, T_0 . Guth and Sher state this clearly, and there is no disagreement here. However, as I stated in my paper, thermal equilibrium cannot be maintained, and it is not clear that the assumption of $T > T_0$ is irrevocable. I shall now consider the possibility of a bounce when these assumptions are changed. I emphasize that I shall discuss this possibility in principle only and that I am not arguing for the extreme fine tuning which is required for a bounce of a closed universe similar to our present universe.

Let us consider a phase transition with the free energy density of the form

$$f(\phi, T) = E_0 \left[1 - \frac{\phi^4}{\sigma_0^4} (1 - 4 \ln \frac{\phi}{\sigma_0}) \right] - G(\phi/T, T) \quad (1)$$

The first term is the Coleman–Weinberg zero-temperature free energy adjusted so that it is equal to zero at $\phi = \sigma_0$. We shall not require the details of the second term describing the temperature effects, except to note that it depends on ϕ/T and asymptotically approaches $T^4 g_\infty/3$ and $T^4 g_0/3$ at low T ($\sigma = \sigma_0$) and high T ($\sigma = 0$), respectively; the constant g s are equal to the $1+g$ factors used by Guth and Sher and by myself (for details see, for example, ref. 3). In equilibrium the Higgs field $\phi = \sigma(T)$, and the value at the minimum of the free energy $\delta f/\delta \phi|_{\sigma(T)} = 0$. In the temperature range $T_1 < T < T_2$, there are two minima, one at $\sigma = 0$ and at $\sigma(T) > 0$, which is almost equal to σ_0 . For $T > T_2$ the non-zero minimum disappears. The critical temperature T_0 of the transition is defined when the two minima are equal:

$$E_0 - G(0, T_0) = f[\sigma(T_0)/T_0, T_0] \quad (2)$$

Given this free energy, one can calculate the pressure, $P = -f$, the specific entropy, $s = -\partial f/\partial T$, and the energy density, $E = Ts + f$:

$$E = E_0(1 - \sigma^4/\sigma_0^4) + 3[G(\sigma/T, T) - E_0(\sigma^4/\sigma_0^4) \ln(\sigma^4/\sigma_0^4)] \quad (3)$$

Note that one can divide the total E and P into 'radiation'-like and 'vacuum'-like parts with $E_r = 3P_r$ equal to the second term and with $E_v = -P_v$ equal to the first term as I had assumed in my paper ($E_v a^4$, with a as the expansion parameter, being equal to the quantity $S^{4/3}$). Such a division, of course, does not change the results. Here I shall use the total energy and pressure.

With this equation and the conservation laws $dE/(E+P) = -3da/a$, it can be shown that if the transition occurs at a constant temperature equal to T_0 , the total entropy of the universe, $\Sigma = 2\pi^2 a^3 s$, remains constant. If the transition occurs at a higher (lower) constant temperature during a contracting (expanding) phase, the entropy will increase: $\Delta \ln \Sigma \approx [|(T_0/T)^4 - 1| \ln g_0/g_\infty]/4$. In either case, a bounce is not allowed. This, however, requires a balance between the phase transition and the contraction (expansion) rates with gravity providing (absorbing) the latent heat of the transition. Such a balance is possible for a second-order phase transition but not for a phase transition described by equation (1) where the universe lingers in a metastable state, $\sigma = 0$, as it cools, or at $\sigma \approx \sigma_0$ as it is heated up. It is such a lingering that gives rise to the inflationary scenario and to a large entropy generation: $\Delta \Sigma \approx \Delta V T_0^3$, where ΔV is the change in the volume of the universe during its inflation.

Similarly, during contraction the universe will remain at the metastable state with $\sigma \approx \sigma_0$ until this minimum of free energy disappears at T_2 . Then, as there is no longer a barrier and

gravity is no longer the governing force for the change of σ towards the lower free energy at $\sigma = 0$, the changes in the Higgs field can occur quickly compared with the contraction time scale. In this case, the conservation equation (which could be written as $dE = -4E, da/a$) reduces to the conservation of energy, $dE = 0$.

Let us now analyse the consequences of this possibility. We first assume that equilibrium is maintained as $\sigma \rightarrow 0$ and $T \rightarrow T_{tr}$. As $\sigma(T_2) \approx \sigma_0$ from equation (3), we find that the energy densities before and after are $3G(\sigma_0/T_2, T_2)$ and $E_0 + 3G(0, T_{tr})$, respectively. Equating these we find

$$G(0, T_{tr}) = G(\sigma_0/T_2, T_2) - E_0/3 \quad (4)$$

The change in entropy density, Δs , as σ changes from σ_0 to zero and T changes from T_2 to T_{tr} , can also be evaluated using equation (3). It can be shown that with the help of the energy conservation relationship (4),

$$T_{tr} \Delta s/4 = G(\sigma_0/T_2, T_2)(1 - T_{tr}/T_2) - E_0/3 \quad (5)$$

Similarly, we can evaluate the quantity $E + 3P$ at $\sigma = 0$, $T = T_{tr}$:

$$(E + 3P)_{tr} = 6[G(\sigma_0/T_2, T_2) - 2E_0/3] \quad (6)$$

One requirement for a bounce is that this quantity must be negative. This and the second law, $\Delta s \geq 0$, demand that

$$(1 - T_{tr}/T_2)^{-1} E_0/3 \leq G(\sigma_0/T_2, T_2) < 2E_0/3 \quad (7)$$

which is satisfied if $T_{tr} \leq T_2/2$. Furthermore, with the help of equation (2) and using the asymptotic relationships for G , it can be shown that this implies $T_{tr} \leq 0.75 T_0$. This disagrees with the conclusions of both Sher³ and the second argument of Guth and Sher¹ against the possibility of a bounce as it shows that the conditions for a bounce are satisfied even when thermal equilibrium is maintained as long as the temperature after the transition is reduced to rather below the critical temperature T_0 from the high metastable value T_2 . I do not believe that Guth and Sher have ruled out this possibility. However, their first argument is valid and as pointed out by Sher (personal communication), the condition (7) places an additional constraint on the details of the transition by requiring that the asymptotic degrees of freedom satisfy the condition $g_\infty/g_0 > 1 + 1.5(T_2/T_0)^4 \approx (T_2/T_{tr})^3 > 8$.

On the other hand, the disappearance of the barrier at T_2 and the required rapid changes may be an indication of the breakdown of equilibrium so that we can no longer describe the state of the universe in terms of temperature or, for example, use the term $G(\sigma, T)$ for energy density of particles. The non-equilibrium particle occupation numbers are required for evaluation of pressure, energy and entropy densities. However, it is obvious that if the Higgs potential increases to E_0 , the energy density in the non-equilibrium distribution will be lower than $3G(\sigma_0/T_2, T_2)$, as in the equilibrium relation (4).

Given such a transition energy density and that during the subsequent contraction E_0 remains constant while energy density in free particles increases as a^{-4} (because of the contraction and the blue shift of energies), it is then possible for the curvature term a^{-2} to catch up, halt the contraction and cause a bounce in a time of $\sim E_0^{-1/2}$.

Without a detailed analysis of the non-equilibrium conditions, it is difficult to determine whether this scenario would work, especially for a bounce of a closed universe as large as our visible universe. A bounce is more likely for lower entropy models which could have existed at the early phase of the universe before the production (through inflation, a previous bounce or otherwise) of an entropy much larger than the entropy of $\sim E_0^{-3/4}$ particles, which is a more natural initial condition than the present value of the entropy of our visible universe.

Received 31 January; accepted 16 February 1983.

1. Guth, A. H. & Sher, M. A. *Nature* **302**, 505–506 (1983).
2. Petrosian, V. *Nature* **298**, 805–808 (1982).
3. Sher, M. A. *Phys. Rev. D* **22**, 2989 (1980).

How stable is our vacuum?

Piet Hut & Martin J. Rees*

Institute for Advanced Study, Princeton, New Jersey 08540, USA

It is possible that the vacuum state we live in is not the absolute lowest one. In many spontaneously broken field theories a local minimum of the effective potential, which can be quite stable, can exist for certain parameter values. The Universe, starting at a high temperature, might have supercooled in such a local minimum. If such a metastable minimum is separated by a high enough barrier from the absolute minimum, the tunneling rate from the 'false' to the 'true' vacuum may be slow enough to not have occurred in one Hubble-spacetime-volume^{1,2}. In that case our vacuum state might suddenly disappear if a bubble of real vacuum formed which was large enough for the bulk energy gain (equal to the product of the volume and the potential drop between false and true vacua) to exceed the surface energy density in its walls (proportional to the barrier potential). Such a bubble would expand at close to the speed of light, with enormous energy release, leaving a large attractive ($\rho = -p < 0$) cosmological constant in the interior, with a geometry close to anti-deSitter space¹. This space-time is singularity-free if a strict vacuum, but any non-zero particle density would cause singularities to develop quickly. Although the persistence of our present vacuum for 10^{10} yr implies that a spontaneous transition via tunnelling is unlikely, we can ask whether a new generation of elementary particle accelerators might trigger such an unfortunate event. We show here that this chance, fortunately, is completely negligible since the region inside our past light cone has already survived some 10^5 cosmic ray collisions at centre of mass energies of 10^{11} GeV and higher.

The most energetic elementary particle reactions occurring on Earth are collisions between very energetic cosmic rays and nucleons in the upper atmosphere. Collisions with centre of mass (c.o.m.) energy exceeding that attainable with present accelerators occur whenever a $\geq 10^6$ GeV cosmic ray hits the Earth's atmosphere. The flux of such particles is such that $\sim 10^8$ reach the Earth every second. This means that any process (at $\sim 10^3$ GeV c.o.m. energy) with a cross-section above the exceedingly low value of $10^{-49} f^{-1} \text{ cm}^2$ has happened at least once in the Earth's history (the factor f is the effective penetration depth of a primary cosmic ray at $\sim 10^6$ GeV, in g cm^{-2}). So there is no chance that existing colliding beam experiments, in a reasonable running time, will trigger any event that would not already have happened in the atmosphere.

But what about still higher energies? Cosmic ray primary energies up to 10^{11} GeV are observed^{3,4}. A collision of such a primary with a nucleon has a c.o.m. energy of $10^{5.5}$ GeV. There may not be any reactions in the atmosphere at energies significantly higher than this, because the spectrum of very high energy cosmic ray primaries is expected to drop steeply above the observed 10^{11} GeV, because of interactions with the 3K background photons (c.o.m. energies > 0.1 GeV, high enough for photopion interactions)^{5,6}. It is not yet clear whether the primaries of giant air showers are single protons or heavier ions, possibly even iron nuclei³. In the latter case the highest observed energy per nucleon will be only 10^9 GeV, or $10^{4.5}$ GeV in the c.o.m. frame of atmospheric collisions. Even $10^{4.5}$ GeV exceeds the highest energies attainable with present-day accelerators by more than an order of magnitude. However, if new types of accelerators would reach this energy, they could trigger reactions more energetic than ever before occurred on Earth.

Despite the rareness of the particles inducing the largest air showers, collisions between two such particles, yielding c.o.m. energies above 10^{11} GeV, may have occurred somewhere in the Universe. We can estimate the collision rate as follows.

Their flux is^{3,4}

$$\phi(E > 10^{11} \text{ GeV}) = 4 \left(\begin{smallmatrix} +2 \\ -1.4 \end{smallmatrix} \right) 10^{-16} \text{ m}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$$

corresponding to a density

$$\begin{aligned} n(E > 10^{11} \text{ GeV}) &= 4\pi c^{-1} \phi(E > 10^{11} \text{ GeV}) \\ &= 1.7 \left(\begin{smallmatrix} +0.8 \\ -0.6 \end{smallmatrix} \right) 10^{-29} \text{ cm}^{-3} \end{aligned}$$

For an order of magnitude estimate we can take all particles to have the same energy $E = 10^{11}$ GeV, with density $n = 10^{-29} \text{ cm}^{-3}$. The appropriate cross-section in this context is, of course, very uncertain, and depends on the parameter values for the barrier height and potential drop. A low barrier with a big drop corresponds to a small critical bubble, easily created in a single collision event. The other extreme—a high barrier with a small potential drop—corresponds to a much larger critical nucleation bubble: here, heavy ion collisions might be more effective. For illustrative purposes, we choose to take a cross-section of the order of the Compton wavelength squared ($\propto E^{-2}$), though this may be too conservative at very high E . For $E \approx 10^{11}$ GeV, we then have $\sigma \approx 10^{-48} \text{ cm}^2$. This leads to a collision rate per particle of $n\sigma c = 3 \times 10^{-67} \text{ s}^{-1}$. The total collision rate per space-time volume is then

$$n^2 \sigma c = 3 \times 10^{-96} \text{ s}^{-1} \text{ cm}^{-3}$$

Our past light cone has a space-time volume of order $c^3 T^4$, where $T \approx 10^{10}$ yr is the Hubble Time. Therefore, the expected number of collisions between ultra-energetic cosmic ray primaries with a c.o.m. energy $E > 10^{11}$ GeV inside our causally connected past is

$$n^2 \sigma c^4 T^4 \approx 10^5$$

This is the number of collisions between particles of comparable energy; an extra contribution arises from collisions between, say, a 10^9 GeV and a 10^{13} GeV particle with the same c.o.m. energy. However, for a power law spectrum of cosmic ray energies this adds only a logarithmic factor (raising the total number of collisions by at most an order of magnitude), and taking into account the 3 K attenuation mentioned above suggest that this factor is even smaller. If the composition of these cosmic rays resembles that at lower energies, some of these collisions will involve iron nuclei.

For higher energies, $E > 10^{12}$ GeV, probably not even one collision has taken place in the history of the observable Universe: even an optimistic extrapolation of the ultra high-energy cosmic ray flux, neglecting any 3 K attenuation, gives^{3,4}

$$\phi(E > E_0) \propto E_0^{-1.5}$$

leading to an expected number of collisions

$$n^2 \sigma c^4 T^4 \approx \left(\frac{E}{10^{12} \text{ GeV}} \right)^{-5}$$

This derivation has assumed a homogeneous distribution of ultrahigh-energy particles. If these particles are clumped on, say, the scales of galaxy clusters then the effective volume for collisions is smaller than $c^3 T^4$ by a factor of $\sim 10^{-2}$. This could reduce the normalization of the energy scale in the previous formula by a factor 2 or 3. But at the place of production of ultrahigh-energy particles the collision probability will have been higher than measured at the Earth, counteracting the effect of the smaller space-time volume available.

Besides the observed high-energy cosmic rays, other objects might exist which could place higher limits on the stability of our vacuum, such as monopoles or even small primordial black holes. The former might induce a transition to a lower vacuum state because of their high mass (of the order of 10^{17} GeV in most models) and small dimensions, while the latter could induce a transition at the moment of final explosion at the end of evaporation by Hawking radiation. However, as long as such objects remain conjectural, neither sets any firm limits on the stability of our vacuum.

* Permanent address: Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK.

We have shown that the most energetic collision in the history of the observable Universe (excluding a small fraction of the first second) had a c.o.m. energy at least of order 10^{11} – 10^{12} GeV. This limit, directly deduced from observations of ultrahigh-energy cosmic ray showers, puts the most stringent limits on the existence of other vacuum states lower than the one we presently inhabit. How seriously the above result constrains a specific spontaneously broken field theory, depends on its parameters such as barrier height and potential drop between true and false vacuum^{1,7}. But at least we can be confident that no particle accelerator in the foreseeable future will pose any threat to our vacuum.

We thank Gary Horowitz and Herbert Neuberger for stimulating discussions. This work was supported in part by the NSF through grant PHY79-19884.

Received 21 December 1982; accepted 16 February 1983.

1. Coleman, S. & De Luccia, F. *Phys. Rev. D* **21**, 3305–3315 (1980).
2. Turner, M. S. & Wilczek, F. *Nature* **298**, 633–634 (1982).
3. Cunningham, G., Lloyd-Evans, J., Pollock, A. M. T., Reid, R. J. O. & Watson, A. A. *Astrophys. J. Lett.* **236**, L71–75 (1980).
4. Bower, A. J., Cunningham, G., Linsley, J., Reid, R. J. O. & Watson, A. A. *J. Phys. G.* (in the press).
5. Greisen, K. *Phys. Rev. Lett.* **16**, 748–750 (1966).
6. Zatsepin, G. T. & Kuzmin, V. A. *Soviet Phys.-JETP Lett.* **4**, 78–80 (1966).
7. Affleck, I. K. & De Luccia, F. *Phys. Rev. D* **20**, 3168–3178 (1979).

Comparison of Earth rotation as inferred from radio interferometric, laser ranging and astrometric observations

D. S. Robertson*, W. E. Carter*, R. J. Eanes†, B. E. Schutz†, B. D. Tapley†, R. W. King‡, R. B. Langley‡§, P. J. Morgan‡ & I. I. Shapiro‡§

* National Geodetic Survey, National Ocean Service, NOAA, Rockville, Maryland 20852, USA

† Department of Aerospace Engineering and Engineering Mechanics, University of Texas at Austin, Austin, Texas 78712, USA

‡ Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

New techniques are being developed to improve the spatial and temporal resolutions achievable in measurements of the rotation vector of the Earth. Such improvements are expected to expose fine details of this rotation and to aid in understanding their physical causes. Although not yet fully developed, two of these techniques are now in routine use: radio interferometry and laser ranging. In the former technique, antennas up to thousands of kilometres apart are used to observe compact extragalactic radio sources. The differences in arrival time of the signals propagating from each source to the antennas, measured by the method of very-long-baseline interferometry (VLBI), are sensitive to variations in each of the components of the Earth's rotation vector. In the latter technique, lasers are used to measure the distance (range) to artificial Earth satellites (satellite laser ranging, SLR) and to optical retroreflectors on the Moon (lunar laser ranging, LLR). To assess the accuracy being achieved in the measurement of Earth rotation it is useful to compare results obtained with these different techniques. National programmes and international cooperation, exemplified by the International Astronomical Union/International Union of Geodesy and Geophysics project MERIT^{1,2}, are encouraging the collection of data suitable for accurate intercomparisons. Here we compare results for a 400-day period from late September 1980 to December 1981, for which data from all three types of measurements are available. We also compare corresponding results from classical astrometric observations which demonstrate the degree of improvement in accuracy afforded by the new techniques.

§ Present addresses: Department of Surveying Engineering, University of New Brunswick, Fredericton, New Brunswick, Canada E3B 5A3 (R.B.L.); Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA (I.I.S.).

The VLBI data, analysed at the National Geodetic Survey (NGS), were obtained in the MERIT short campaign and from the NGS's project POLARIS³. The MERIT data, collected during two 1-week observing sessions in September and October 1980, resulted from observations of up to five antennas operating simultaneously, three in the USA and two in Europe. The POLARIS observations were made at 2-week intervals from November 1980 to June 1981, and at 1-week intervals thereafter. Each POLARIS session spanned ~24 h, with most involving only the Haystack or Westford Observatory in Massachusetts and the Harvard Radio Astronomy Station in Texas. The Onsala Space Observatory in Sweden also participated in several of the sessions. VLBI data from two-antenna (single-baseline) observations can determine only two linear combinations of the three parameters required to describe the location of the rotation pole on the Earth's surface (X and Y) and the angle of rotation about this pole (universal time, UT1). Our single-baseline observations were most sensitive to the X value for the pole and to UT1. Estimates of Y , which could be determined usefully only for the six POLARIS sessions involving Onsala and for the 2 weeks of the MERIT session, will not be discussed here. In the reduction of the single-baseline VLBI data, Y was obtained by interpolation from the 'Circular D' values of pole position published by the Bureau International de l'Heure (BIH). A representative error of 10 m arcs in the Circular D value of Y introduces corresponding errors of 3 m arcs and 0.2 ms of time in the deduced values of X and UT1, respectively. These errors are small enough for our purposes to obviate the need for a more accurate set of Y values.

The SLR data, analysed at the University of Texas, and obtained from multi-station observations of the LAGEOS satellite, were collected as part of the National Aeronautics and Space Administration's (NASA's) Geodynamics Program^{4,5}. The satellite's orbit is a useful reference for interpreting effects of polar motion on the tracking stations. The stations' positions relative to the orbit change with time due to polar motion as well as to gravitational and other effects on the orbit. The current accuracies of orbit computation allow separation of these various effects. As a consequence, the SLR observations have been used since the MERIT short campaign to derive a set of values for the X and Y components of the pole position from each successive (non-overlapping) 5-day interval. To obtain values at the epochs corresponding to the VLBI determinations, and also to reduce the statistical scatter, the SLR values were smoothed by convolution with a gaussian function having a full-width at half-maximum (FWHM) of 10 days. The root-mean-square (r.m.s.) scatter of the 5-day values about the smoothed curve are 4 and 3 m arcs for X and Y , respectively.

The LLR data, analysed at the Massachusetts Institute of Technology, were obtained at the University of Texas McDonald Observatory, also as part of NASA's Geodynamics Program^{6,7}. LLR observations from a single station determine only two linear combinations of the parameters X , Y and UT1. For the McDonald Observatory data, there is far higher sensitivity to UT1 and to Y than to X . The LLR estimates of Y were restricted by geometry to a temporal resolution of the order of 1 month, too long to be of interest here. We calculated UT1 from the appropriate linear combination using interpolated SLR values of X . A representative error of 5 m arcs in X as determined from SLR data causes an error in the LLR UT1 estimates of ~0.2 ms of time.

LLR estimates of UT1 were made for each day for which adequate observational data existed—about 50 days over the 400-day span being discussed. The values from individual days were smoothed by convolution with a gaussian function with a FWHM of 8 days for other purposes⁷, but they also serve for intercomparison in that values are thereby available for all epochs. The r.m.s. scatter of the individual-day values about the smoothed curve is 0.5 ms of time.

Figure 1 displays the VLBI, SLR and classical astrometric⁸ determinations of X shown as differences from reference values determined from heavily smoothed astrometric measurements.

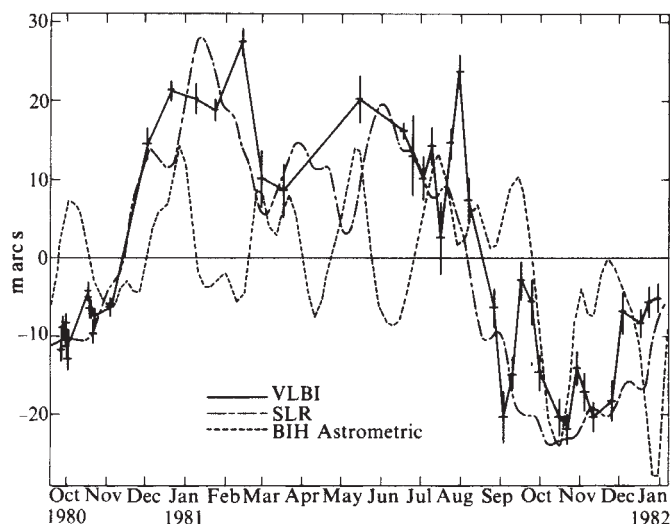


Fig. 1 VLBI, SLR and BIH astrometric⁸ determinations of X , shown as differences from the values determined from heavily smoothed astrometric measurements (see text), after removal of the mean differences between the various sets of values. The vertical bar on each VLBI point indicates the statistical standard error of the estimate (no allowance for possible systematic errors). The SLR and BIH astrometric values have been convolved with a gaussian function with FWHM of 10 days. The lines connecting the VLBI points are intended only to guide the eye, not to imply the actual pole track. (1 m arcs is equal to a displacement of ~ 3 cm at 1 Earth radius.)

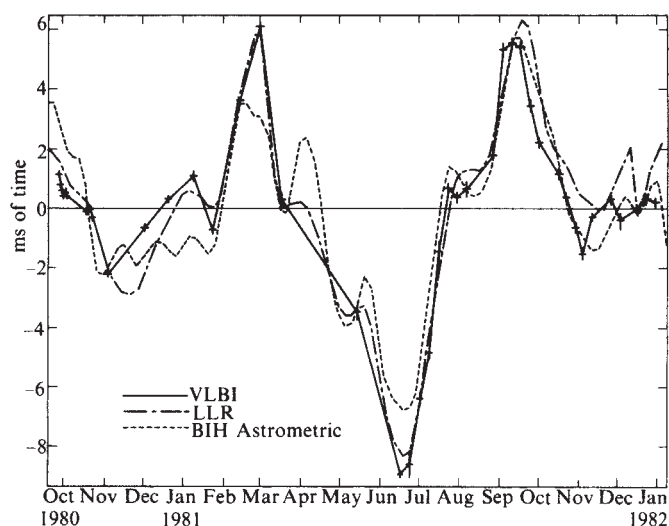


Fig. 2 VLBI, LLR and BIH astrometric⁸ determinations of UT1 (see text), plotted as differences from heavily smoothed astrometric determinations, after removal of the mean differences between the various sets of values. As in Fig. 1, the error bars on the VLBI data indicate the statistical standard errors (no allowance for possible systematic errors), and the lines connecting the VLBI points are intended only to guide the eye. (1 ms of time is equivalent to a displacement of nearly 50 cm at the Earth's Equator.)

The astrometric values shown as a broken curve in the figure were obtained by smoothing the original values by convolution with a gaussian function with FWHM of 10 days, whereas the reference values were obtained by convolution with a gaussian function with a FWHM of 40 days. As there is an arbitrary origin in the pole position, determinations from any of these techniques, the mean difference between each of these curves and the reference values has been removed.

Figure 1 shows that the VLBI and SLR values often follow similar trends. However, there are many seemingly significant differences, especially in July and September 1981. The relative contributions of smoothing and of errors in the VLBI and SLR estimates to these differences are not known. The r.m.s. difference between the VLBI and SLR estimates is 6 m arcs (18 cm) when evaluated for the epochs of the POLARIS VLBI observations and for the first and last days only of each week of MERIT observations so as to include only nearly independent SLR values in the comparison. The corresponding differences between the classical set of values and the VLBI and SLR sets were each 14 m arcs.

Although the values of pole position published regularly by the BIH in its Circular D include a nearly 50% contribution from the SLR results and are more heavily smoothed in a practically one-sided manner to allow timely distribution⁹, Circular D has such wide usage that it is of interest to compare its values with the sets of VLBI and SLR numbers presented here. Both r.m.s. differences, calculated as for the corresponding comparisons given above, are 9 m arcs.

Figure 2 displays the VLBI, LLR and classical astrometric determinations of UT1, shown as differences from reference values determined from heavily smoothed astrometric measurements as in Fig. 1. Also as in Fig. 1, the mean difference between each of these curves and the reference values has been removed. At times the trends in the VLBI and LLR values match fairly closely, although there are several significant discrepancies as yet unexplained, especially in the interval between November 1980 and January 1981. The r.m.s. difference between the VLBI and LLR values of UT1, evaluated as for the values of X , is 0.9 ms of time, equivalent to ~ 40 cm at the Equator. The

corresponding r.m.s. differences between the VLBI and LLR sets and the classical set are both 1.2 ms of time (55 cm). The r.m.s. differences between the VLBI and LLR values of UT1 and those of Circular D are 1.6 and 1.4 ms of time, respectively. These last comparisons are doubtless adversely affected by the heavier smoothing of the Circular D values.

Our results indicate that VLBI and SLR, at their present stages of development, yield standard errors under 20 cm in the determinations of X , about twofold smaller than obtained from classical measurements, and that VLBI and LLR yield determinations of UT1 with standard errors less than 40 cm, somewhat smaller than that of the corresponding determinations from classical observations.

There are several ways in which these types of intercomparisons can be improved. The density of the VLBI observations can be increased: Polaris observations are expected to continue at weekly intervals until the beginning of the Merit main campaign in September 1983, at which time the frequency should increase to at least once each 5 days and the observations should include at least three antennas. The spatial and temporal resolutions achievable regularly from the SLR observations will increase gradually as the ranging accuracies improve, as the number and distribution of stations increase worldwide, and as the model for the LAGEOS orbit determination improves. The density of the LLR measurements will increase markedly only if additional observatories participate.

The ultimate goal, as mentioned, is not merely to measure variations in Earth rotation more accurately, but to determine their underlying geophysical causes. Careful assessment of the accuracy of the measurements is critical to achieving this goal. Appropriate intercomparisons should provide a good measure of these accuracies.

We thank the staffs of the participating observatories as well as the NASA Geodynamics program for their assistance. The University of Texas experimenters were supported by NASA contracts NAS5-25991 and NAS5-26955. The MIT researchers were supported by NASA contract NAS5-25833 and Air Force Geophysics Laboratory contract F19628-82-K-0002. P.J.M. was supported by the Australian Division of National Mapping.

Received 2 December 1982; accepted 7 February 1983.

1. Wilkins, G. A. & Feissel, M. (eds) *Project MERIT: Report on the Short Campaign and Grasse Workshop with Observations and Results on Earth-Rotation During 1980 August-October* (Royal Greenwich Observatory, Herstmonceux, 1982).
2. Wilkins, G. A. (ed.) *Project MERIT: A Review of the Techniques to be Used During Project MERIT to Monitor the Rotation of the Earth* (Royal Greenwich Observatory, Herstmonceux, 1980).
3. Robertson, D. S. & Carter, W. E. in *High Precision Earth Rotation and Earth-Moon Dynamics* (ed. Calame, O.) 97-122 (Reidel, Dordrecht, 1982).
4. Tapley, B. D., Schutz, B. E. & Eanes, R. J. *Proc. 4th IAG Laser Ranging Workshop*, 523-567 (University of Texas, Austin, 1981).
5. Schutz, B. E., Tapley, B. D. & Eanes, R. J. in *A. Rep. 1980, D27-D36* (Bureau International de l'Heure, Paris, 1981).
6. Langley, R. B., King, R. W., & Shapiro, I. I. *J. geophys. Res.* **86**, 11913-11918 (1981).
7. Langley, R. B., King, R. W., Morgan, P. J. & Shapiro, I. I. in *High Precision Earth Rotation and Earth-Moon Dynamics* (ed. Calame, O.) 25-30 (Reidel, Dordrecht, 1982).
8. Feissel, M. in *A. Rep. 1980, 1981* (Bureau International de l'Heure, Paris, 1981, 1982).
9. *A. Rep. 1981* (Bureau International de l'Heure, Paris, 1982).

Condensation nuclei events at 30 km and possible influences of solar cosmic rays

D. J. Hofmann & J. M. Rosen

Department of Physics and Astronomy, University of Wyoming, Laramie, Wyoming 82071, USA

Two recent observations have provided the basis for study of a relationship between solar activity and the formation of small particles in the Earth's atmosphere: the discovery of annual increases of condensation nuclei (CN) at 30 km (refs 1-3) and the detection of sulphuric acid molecules in large negative ion clusters in the 25-35-km altitude region⁴⁻⁷. These observations have now led us to formulate and test a model wherein CN are formed in a 'polar cloud chamber' supersaturated with sulphuric acid vapour and triggered by ionization associated with solar flare cosmic radiation. We conclude that such a model provides a potential explanation of the observations.

Vertical profiles of CN ($r > 0.01 \mu\text{m}$) have been measured routinely from Laramie, Wyoming (41°N), with optical particle counters since 1973. These observations, and descriptions of the instrumentation, have been presented previously^{2,3,8-10}. Observations of annual CN 'events' since 1979 are shown in Fig. 1. Pre-event profiles are indicated by dotted curves. Some of the CN layers below 25 km altitude in Fig. 1 are due to the volcanic activity of Mt St Helens in 1980, Alaid in 1981 and El Chichón in 1982 and will not be discussed here. In the region above 25 km, the nature of the CN event manifests itself through an increase in the mixing ratio (number of CN per mg of ambient air) by about an order of magnitude during the period February-April. Trajectory analysis indicates that the events probably originate in the polar region¹⁰.

Several size and volatility experiments were performed on the event layer in 1982 to determine its character^{3,10}. They indicated that the CN event when first observed, consisted of newly formed, small sulphuric acid droplets. Later in the year the particles are larger due to coagulation and growth, indicating that they are probably not being formed continuously at a reduced rate throughout the year.

Figure 1 indicates that the quiescent level of CN at 25-30 km has increased from about 100 mg^{-1} before 1981 to a value of about 200 mg^{-1} after the 1981 event, suggesting that the recent intensity of the events has been sufficient to alter the equilibrium level of CN at 30 km. As CN are precursors of larger optically active aerosol, this observation may be important in the subsequent formation of a new scattering layer at these heights.

The second recent observation, important to this analysis, involves ion mass spectrometer measurements at high altitude over southern France and the subsequent identification of sulphuric acid molecules in large negative ion clusters by Arnold and co-workers⁴, an identification verified by independent measurements⁷. Extended measurements in September 1980, by this technique, allowed the vertical distribution of H_2SO_4

vapour between 23 and 38 km to be determined and indicated a concentration of 10^6 - 10^7 cm^{-3} at 35 km (ref. 11). It was further shown that the measured vapour pressure was nearly equal to the saturation vapour pressure between approximately 30 and 35 km. These measurements of vapour mass are in reasonable agreement with our observed sulphate aerosol mass at about 33 km, which is what would be expected from simple considerations³, and also suggest the possible importance of ion nucleation of sulphuric acid droplets².

Arnold has recently re-examined the role of ion nucleation in the formation of sulphuric acid aerosol in the stratosphere and found it to be effective in certain conditions such as sudden stratospheric warming and subsequent sudden cooling to increase the level of H_2SO_4 supersaturation⁶. Several other mechanisms capable of effectively increasing the level of H_2SO_4 saturation have been proposed, including increased levels of H_2SO_4 vapour present in polar regions as might occur through sunlight oxidation processes in spring², volcanic eruptions and subsequent diffusion to the 30-km region and transport poleward³, or gradual winter cooling in polar regions³. The latter 'polar cloud chamber' seems to be a viable mechanism for obtaining sufficient supersaturation to create the CN event except that winter cooling in north polar regions begins in September while the CN events have not been conclusively observed before early February. Thus, some additional triggering mechanism, such as stratospheric warming-cooling episodes to produce additional supersaturation or high altitude ionizing events to increase the ion concentration, may be necessary for consistency with the data. Similarly, the photochemical production of H_2SO_4 , when sunlight returns to the polar regions after winter darkness, is probably not in itself fast enough to produce the observed effects. Several of these mechanisms, occurring coincidentally, may be required.

As the Sun directly or indirectly controls all ionizing radiations received by the upper atmosphere, it is logical that if ion nucleation is important in forming high altitude CN, there may be a cause-effect relationship between solar activity and the creation of condensation nuclei in the upper atmosphere. To investigate further the effect of solar activity on the particle nucleation rate, it is necessary to examine the relationship between the ion production rate and the nucleation rate.

The production of stable neutral clusters of H_2SO_4 and H_2O molecules depends on the rate of growth through association of such molecules and the rate of recombination of the growing negative ion with a positive ion before a size can be reached at which the cluster is stable even after neutralization⁶. The number of H_2SO_4 molecules required to reach critical size is not well known but conservatively lies between 3 and 10. Critical embryos then grow by further association of H_2SO_4 and H_2O and self-coagulation to CN-sized ($r \sim 0.01 \mu\text{m}$) particles.

Simple consideration of the competition between H_2SO_4 association and ion recombination at steady state⁶ gives the approximate relationship between nucleation rate J ($\text{cm}^{-3} \text{ s}^{-1}$) and ion production rate Q ($\text{cm}^{-3} \text{ s}^{-1}$) as

$$J = Q \left(1 + \frac{t_a}{t_r} \right)^{-N_c} \quad (1)$$

where $t_a = [kn(\text{H}_2\text{SO}_4)]^{-1}$ is the characteristic time for H_2SO_4 association, where k is the association rate coefficient ($\sim 10^{-9} \text{ cm}^3 \text{ s}^{-1}$) and $n(\text{H}_2\text{SO}_4)$ is the molecular concentration of H_2SO_4 . $t_r = (\alpha n_{\pm})^{-1}$ is the characteristic time for ion-ion recombination, where α is the ion recombination coefficient ($\sim 10^{-7} \text{ cm}^3 \text{ s}^{-1}$) and $n_{\pm}$ the positive or negative (assumed equal) ion concentration ($\sim 5 \times 10^3 \text{ cm}^{-3}$). N_c is the number of H_2SO_4 molecules in a critically sized embryo.

Since Q and $n_{\pm}$ are related at steady state by $n_{\pm} = (Q/\alpha)^{1/2}$, J may be written

$$J = Q \left(1 + \frac{(\alpha Q)^{1/2}}{kn(\text{H}_2\text{SO}_4)} \right)^{-N_c} \quad (2)$$

in terms of Q , $n(\text{H}_2\text{SO}_4)$ and N_c , which are the basic variables

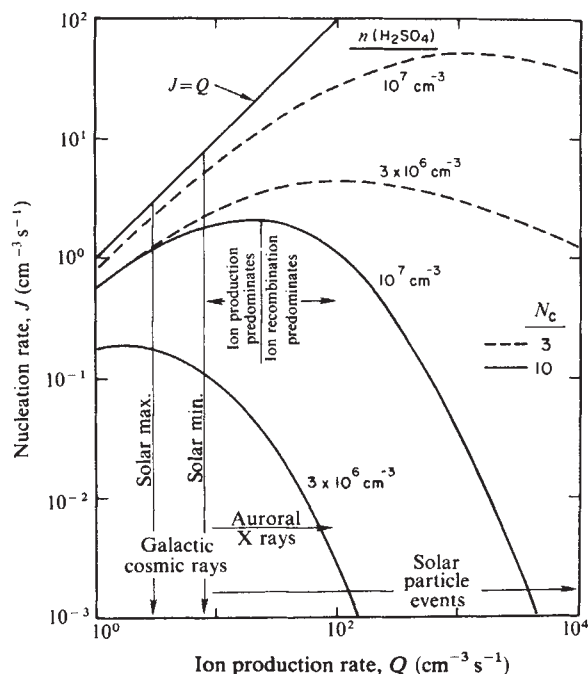


Fig. 2 Ion nucleation rate versus ion production rate for several values of sulphuric acid vapour concentration and the number of sulphuric acid molecules constituting a critical embryo (N_c). The ion production rate for various ionizing radiations are indicated.

conclude that ion nucleation is a viable CN event source and that an $N_c \sim 3$ and an initial critical embryo concentration of about 10^4 – 10^5 cm^{-3} are reasonable estimates of these parameters. Further refinement of our estimates must await a complete model of nucleation, growth, coagulation and diffusion, taking into account the changing size distribution with some knowledge of the size of the source region.

These considerations and the fact that the sulphuric acid concentration is presently quite high due to recent volcanic activity then lead us to a best current estimate of J versus Q as the upper dashed curve in Fig. 2. We would thus expect J to increase very nearly linearly with Q for the transition from solar maximum (galactic cosmic ray minimum) to solar minimum (galactic cosmic ray maximum), for most auroral X-ray–electron precipitation events and for small to average solar flare particle events. Only the largest such events (the giant solar flare particle event of 4 August 1972 has been estimated to have produced ions at a rate of $\sim 10^5$ $\text{cm}^{-3} \text{ s}^{-1}$ at 30 km) could effectively reduce the nucleation rate.

We have investigated whether variations in solar activity have had a general influence on CN production and concluded that there may exist a long-term correlation between CN concentration at 20 km and solar activity, with higher concentrations at solar maximum². This is the opposite of what one would expect if galactic cosmic rays were the dominant ionization mechanism. In reality, insufficient data are available to study effectively a long-term variation due to variations in galactic cosmic rays¹⁰. With the meagre data base at hand, we can determine if the intensity of CN events has varied with, for example, solar flare activity. If the events were related to ionizing episodes such as solar flare particle events, then we would expect their intensity and duration to be proportional to such activity.

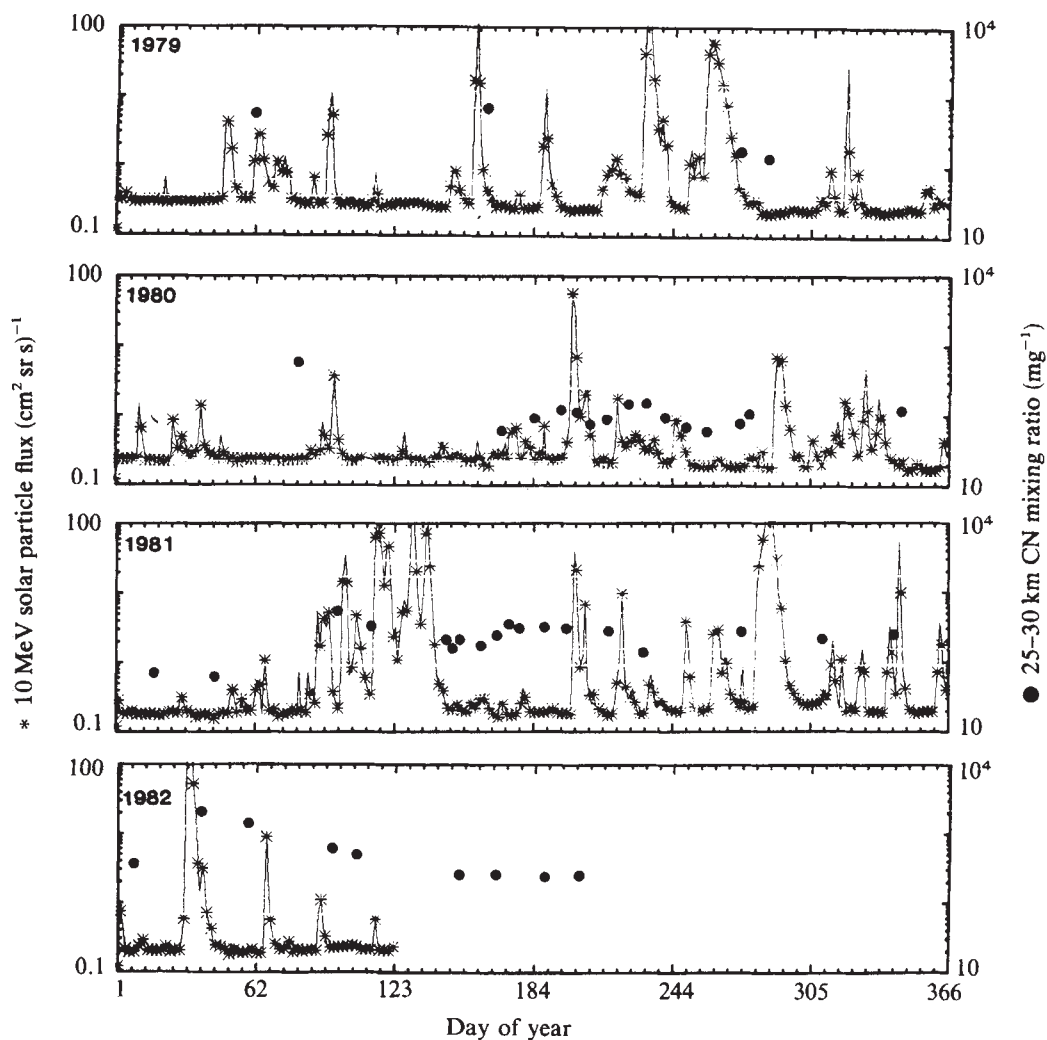


Fig. 3 The particle flux of $E \geq 10$ MeV solar flare events and the 25–30 km condensation nuclei mixing ratio for the past 4 yr.

While most of the solar proton events which occurred over the past 4 yr did not produce many particles energetic enough to penetrate directly to 30 km, the ionizing effects of protons in the 1–17-MeV energy range can be readily felt at this altitude through the penetrating γ rays produced in nuclear interactions at the top of the atmosphere¹³. In addition, major auroral activity and removal of geomagnetic cutoff rigidities at mid-latitudes, during the geomagnetic storm following large solar flares, are further causes of enhanced atmospheric ionization at this time. Whereas 1980 was low in frequency of proton events, 1979 and especially 1981 were quite high, in general agreement with the magnitude of the CN event observations. The CN event mixing ratio and the >10-MeV solar particle flux¹⁴ are plotted against time for 1979–82 in Fig. 3. Although the CN event at first observation seems to correlate with a solar particle event in 1979, 1981 and 1982, it does not do so in 1980. There seems to be little correlation during the latter half of the year, possibly due to a lack of sufficient H_2SO_4 supersaturation during these warmer periods. Unfortunately, the CN measurements are not continuous and further complications arise from transport effects; however, the data do not preclude the possibility that solar flare related ionizing particle events serve as a triggering mechanism when the H_2SO_4 vapour is supersaturated. This process is similar to that governing the operation of a Wilson cosmic ray cloud chamber except that the supersaturation would be obtained through cooling rather than by expansion.

In conclusion, a model of CN production through ion nucleation in an H_2SO_4 supersaturated polar cloud chamber, possibly triggered by solar flare related ionization, seems to represent a mechanism which explains most features of the CN events. The new CN develop only during periods of H_2SO_4 supersaturation at 30 km, in a region accessible to solar ionizing radiation, that is, in the winter polar regions. They are small but numerous enough to survive transport to mid-latitude. The events seem to intensify during volcanically active periods, when more H_2SO_4 vapour is available, and in a general sense, are more frequent during periods of maximum solar activity, the latter a feature reported earlier². It is quite clear that in sufficient abundance, as observed recently, the event CN can seed the 20-km sulphate layer and as such may be important in climatic considerations. The recent (4 April 1982) eruption of El Chichón in Mexico, which has severely perturbed stratospheric aerosol levels to at least 30 km, has apparently created a high level of H_2SO_4 supersaturation ($>10^7$ molecules cm^{-3})¹⁵ in this region and set the stage for the possible occurrence of very large CN events in the next few years.

We thank F. Arnold for useful discussion. This research was supported jointly by NASA, the NSF and the Office of Naval Research.

Note added in proof: On 28 January 1983, at 29 km altitude, we observed an abrupt increase in what appear to be small sulphuric acid droplets of more than 600 cm^{-3} which persisted to our maximum altitude of 33 km. This is at least 15 times larger than the 1982 event and is almost certainly due to El Chichón injected vapours which have been transported to the polar regions in winter either directly or through sulphuric acid aerosol and have subsequently evaporated.

Received 17 November 1982; accepted 1 February 1983.

- Rosen, J. M. & Hofmann, D. J. *Rep. AP-8* (Dept of Physics & Astronomy, Univ. Wyoming, December, 1981).
- Hofmann, D. J. & Rosen, J. M. *Nature* **297**, 120–125 (1982).
- Rosen, J. M. & Hofmann, D. J. *J. geophys. Res.* (in the press).
- Arnold, F. & Fabian, R. *Nature* **283**, 55–57 (1980).
- Arnold, F. & Henschen, G. *Geophys. Res. Lett.* **8**, 83–86 (1981).
- Arnold, F. *Nature* **299**, 134–137 (1982).
- Aris, E., Nevejans, D., Frederick, P. & Ingels, J. *Geophys. Res. Lett.* **8**, 121–124 (1981).
- Rosen, J. M. & Hofmann, D. J. *J. appl. Met.* **16**, 56–62 (1979).
- Rosen, J. M., Hofmann, D. J. & Kaselau, K. H. *J. appl. Met.* **17**, 1737–1740 (1978).
- Hofmann, D. J. & Rosen, J. M. *Rep. AP-73* (Dept. of Physics and Astronomy, Univ. Wyoming, August, 1982).
- Viggiano, A. A. & Arnold, F. *Geophys. Res. Lett.* **8**, 583–586 (1981).
- Yue, G. K. & Hamill, P. J. *Aerosol Sci.* **10**, 609–614 (1979).
- Hofmann, D. J. & Winckler, J. R. *J. geophys. Res.* **68**, 2067–2098 (1963).
- Armstrong, T. P., Brungard, C. & Meyer, J. *Weather and Climate Responses to Solar Variations* (ed. McCormac, B. M.) (Colorado Associated University Press, 1982).
- Hofmann, D. J. & Rosen, J. M. *Geophys. Res. Lett.* (in the press).

The M2 oceanic tide recovered from Seasat altimetry in the Indian Ocean

Pierre Mazzega

CNES/GRGS, 18 Av. E. Belin, 31400 Toulouse, France

Seasat altimeter data, which provide measurements of the instantaneous ocean level, are used here to create a model of the M2 oceanic tide in the Indian Ocean. This approach avoids the assumptions introduced in the hydrodynamical tide models and takes advantage of a regular distribution of the data over the whole ocean. The hydrodynamical and empirical models that have been proposed for this ocean^{1–11} exhibit strong discrepancies, particularly in the deep ocean so that the use of altimeter data is of particular interest. Attempts to recover the tides in limited areas of the Atlantic and Pacific oceans from altimeter data have been made previously^{12–17}. We investigate a new method to obtain the amplitude and phase of M2 in a large region. The validity of the model is checked using synthetic data and by comparison with tide gauge data, and seems to be reliable in the deep ocean.

Each altimeter measurement gives the instantaneous sea level with respect to a reference ellipsoid. It includes mainly the geoid height, the quasi-permanent mean circulation, the tides and the radial component of the orbit error. It has been shown that, before any tidal analysis, it is necessary to remove the geoid¹⁸. We have found that a simple subtraction of a published geoid model is not sufficiently accurate¹⁹, particularly at the short wavelengths. Consequently, we estimate the mean ocean level directly from a mean of altimetric data at locations where enough data are available. For this purpose, we have selected the data of the last 25 days of the Seasat experiment, 15 September–8 October 1978, when Seasat had a 3-day repeating orbit, so that eight satellite passes are available over each track, and 16 values at each cross-over (Fig. 1). Data are processed as follows: outliers and data over land or sea ice are eliminated. Along each profile and for each cross-over, we select and average the data in the vicinity of the cross-over. The mean of the 16 values which depicts the geoid and the quasi-steady oceanic effects is removed from the raw data. The residuals, which are mainly due to the orbit error¹³, follow a gaussian distribution with a 1.07 m r.m.s. (Fig. 2). The corrections for the atmospheric pressure effect²⁰ and for the sea state bias effect²¹ (–5% of the significant wave height) do not reduce significantly the data scattering. Hereafter we assume that the orbit error is random with respect to the tide; this point will be discussed later.

The orbit repeat period (3 days and 13 min for Seasat) and the Earth rotation period combine to alias tidal components into long periods when observed by altimeter (D. E. Cartwright, personal communication). Table 1 shows that with the 25 days of available data, only the M2, N2, O1 partial tides can be extracted. The other components are nearly frozen with respect to this time scale so they are equivalent to the steady oceanic effects. However, N2 can be neglected because of its low amplitude. The closeness of the apparent period of M2 and O1 makes it difficult to resolve the two components. We deal only with the main component M2 and then verify with synthetic data that the other tidal components do not alias the M2 solution.

The altimeter data give information about the instantaneous geocentric tide ζ^g which includes the solid earth tide ζ^e , the oceanic tide ζ^o , and the yielding of the earth ζ^1 in response to the oceanic tide load:

$$\zeta^g = \zeta^e + \zeta^o + \zeta^1$$

where ζ^1 is related to the oceanic tide ζ^o by²²

$$\zeta^1 = -0.0667 \zeta^o$$

The altimeter data are corrected for the solid earth tide²³.

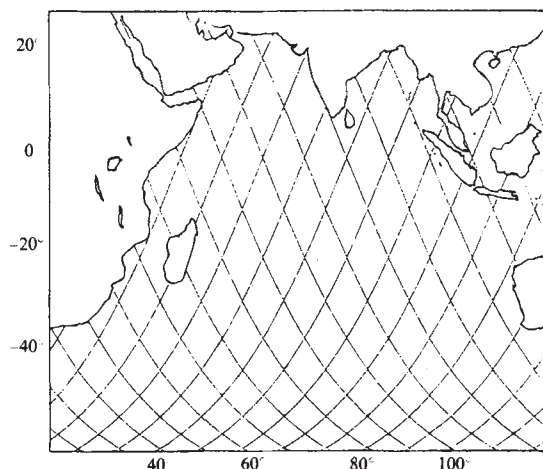


Fig. 1 Coverage of the Seasat altimeter data of the repeating orbit over the Indian Ocean.

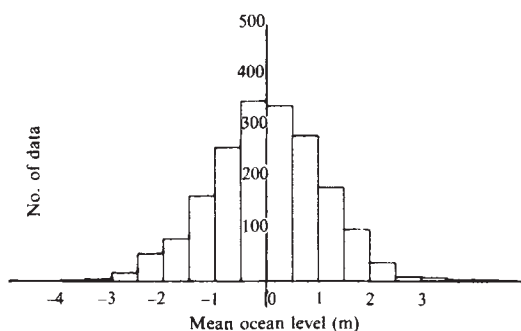


Fig. 2 Histogram of the residuals (altimeter data-mean ocean level). Mean = 0.01 m, r.m.s. = 1.07 m.

We perform a two-dimensional tidal analysis of the altimeter values to get a smooth solution for the amplitude and for the phase of the M2 component over the entire Indian Ocean. The data grid has meshes of about $8^\circ \times 8^\circ$ near the Equator so that the retrieval of wavelengths shorter than 16° is impossible.

If the orbit error can be neglected, we may represent an altimetric value $H(\theta, \lambda, t)$ as⁸:

$$H(\theta, \lambda, t) = U(\theta, \lambda) \cos(\Omega t + \phi) + V(\theta, \lambda) \sin(\Omega t + \phi) + dN(\theta, \lambda) \quad (1)$$

where (θ, λ) are the geographical coordinates, t is the universal time, Ω is the M2 angular frequency and ϕ is the M2 astronomical argument. dN is introduced to account for an eventual static residual. The amplitude $A(\theta, \lambda)$ and the phase $P(\theta, \lambda)$ of the M2 tide are related to U and V by $U = A \cos P$ and $V = A \sin P$. To obtain spatially coherent solutions, each of the unknown functions U , V , dN , hereafter called F^i , is expanded into a series of surface spherical harmonics:

$$F^i(\theta, \lambda) = \sum_{l=0}^{l_{\max}} \sum_{m=0}^l \{a_{lm}^i \cos(m\lambda) + b_{lm}^i \sin(m\lambda)\} P_{lm}(\sin \theta) \quad (2)$$

where the a_{lm}^i and the b_{lm}^i are constant coefficients and the P_{lm} are associated Legendre's functions.

The expansions of the U , V and dN are introduced in equation (1) and the coefficients are determined by a standard least squares regression analysis. The number of unknowns $3(l_{\max} + 1)^2$ increases steeply with increasing $l_{\max}$. The co-range (equi-amplitude) and co-tidal (equi-phase) maps of the M2 oceanic tide alone (without the loading effect) deduced from the U and V solutions for $l_{\max} = 3$ are given in Fig. 3. This

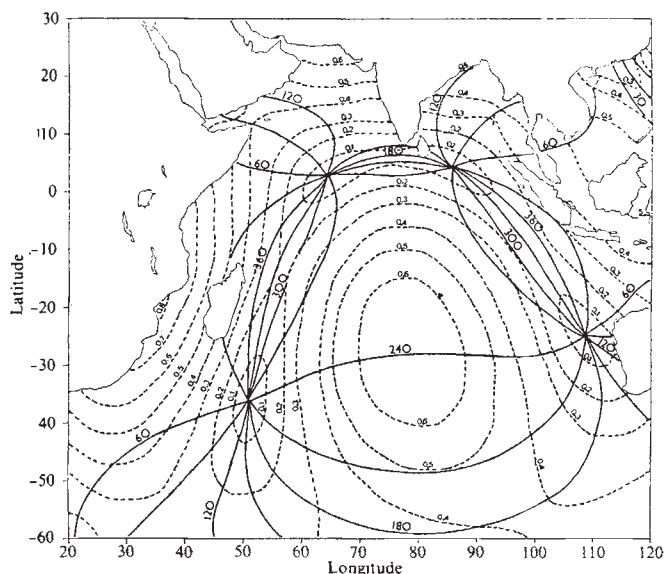


Fig. 3 Map of the M2 oceanic tide solution: cotidal lines in degrees (—); corange lines in metres (----).

Table 1 Amplitudes, periods and aliased periods of the leading partial tides

Partial tides	Amplitude relative to M2	Period (h)	Aliased period (days)
M2	1	12.42	16
S2	0.4656	12.00	170
N2	0.1915	12.66	10
K2	0.1267	11.97	91
K1	0.5842	23.93	170
O1	0.4148	25.82	14
P1	0.1933	24.07	∞
Q1	0.0795	26.87	10

model explains only 6% of the variance of the data: this low value is related to the very low signal-to-noise ratio due to the orbit error. For $l_{\max} = 4$, the explained variance does not increase significantly (6.5%). For greater values of $l_{\max}$, the solutions exhibit instabilities probably induced by the noise of the data and by the difficulty of modelling, with spherical harmonics, functions defined over a relatively small area.

To test the validity of the solution, we have applied the same analysis procedure to synthetic data. We computed 16 instantaneous values relevant to the times of the real measurements, using the Schwiderski model of the nine leading tides²⁴, namely M2 (Fig. 4a), S2, N2, K2, O1, P1, Q1 and M_t at each cross-over. The M2 solution obtained using the above method with $l_{\max} = 3$ (Fig. 4b) explains 87% of the variance of the M2 signal and the main features are well recovered. Near the shores the M2 recovery is not successful everywhere, but we have checked that, with noise-free data, an increase of $l_{\max}$ improves the solution (94% of explained variance with $l_{\max} = 4$). Considering the small difference between the M2 and O1 aliased periods (Table 1), the method we applied appears to be very selective in the frequency domain.

As indicated previously, the influence of the orbit error over the solution must be carefully analysed¹³⁻¹⁸. We therefore added to the synthetic data a simulated orbit error. The spectrum of this error includes two peaks with a gaussian distribution, a dominant one at the 14.3 cycles day⁻¹ (ref. 25) with a 160-cm amplitude and 0.25 cycles day⁻¹ r.m.s., the other one at the semi-diurnal frequency with a 5-cm amplitude and 0.1 cycles day⁻¹ r.m.s. To this point-by-point generated error, we added along each track an error due to a random tilt with mean zero and standard deviation = 2×10^{-3} cm km⁻¹. With this orbit error

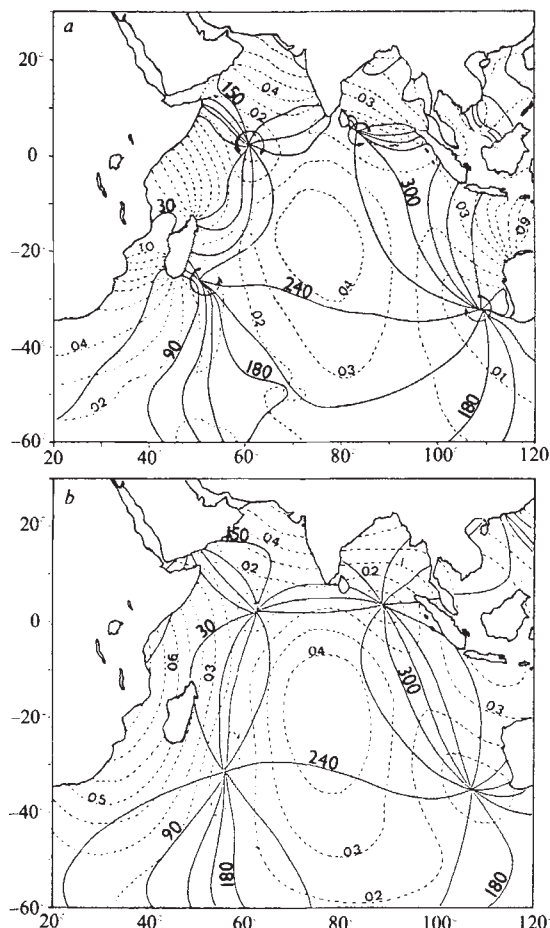


Table 1 Isotopic measurements of francolite samples from different phosphate-deposit areas

Sample	Area	$\delta^{34}\text{S}$ (‰)	$\delta^{13}\text{C}$ (‰)	Age	$\delta^{34}\text{S}$ open ocean sulphate evaporites of same age*
310A	Agulhas Bank	+20.8	-0.70	Miocene	+20.5-+23.0
BP2	Blake Plateau	+20.3	-0.23	Miocene	+20.5-+23.0
AHF14402	Californian Borderland	+20.3	-1.17	Miocene	+20.5-+23.0
GMI-CX-79	Chatham Rise	+22.3	-0.81	Miocene	+20.5-+23.0
1022(1)	Moroccan Shelf	+22.3	-3.79	Upper Cretaceous	+14.0-+20.0
148	Moroccan Shelf	+22.7	—	Eocene	+18.0-+22.0
USGS23	Phosphoria Formation, USA	+15.3	-8.02	Upper Permian	+9.5-+12.5
PD-19-33(L)	Peru/Chile Shelf	+10.7	+0.27	Recent	+20.3-+22.2
3454	Namibian Shelf (pelletal)	+10.8	-3.31	Miocene	+20.5-+23.0
SA/V7(B)/79	Onshore South African (pelletal)	+10.3	-4.50	Miocene	+20.5-+23.0
SA/V3(C)/79	Onshore South African (pelletal)	+14.2	-9.52	Miocene	+20.5-+23.0
5799	Offshore Namibian Concretion	+19.3	-8.91	Recent	+20.3-+22.2

$\delta^{34}\text{S}$ with respect to Cañon Diablo, $\pm 0.17\%$ standard error at the 95% confidence level. $\delta^{13}\text{C}$ with respect to PDB, $\pm 0.03\%$ standard error at the 95% confidence level.

* Data from ref. 13.

of sulphate occurs (by the oxidation of organic matter). Authigenic francolite, forming in these conditions in anoxic muds would therefore be expected to contain an appreciable component of organically derived carbon which would make it isotopically light⁴. In contrast, replacement francolites inherit similar $\delta^{13}\text{C}$ compositions to their precursor carbonate³⁻⁶ which swamps the small bicarbonate reservoir on dissolution. This isotopic characterization relates to the genetic environment of primary phosphate and must be used with caution in interpreting the origin of phosphorite deposits which may consist of reworked material.

Although there has been a fluctuation in the $\delta^{34}\text{S}$ value of seawater through geological time, the relatively long residence time ($\sim 2 \times 10^7$ yr) of dissolved sulphate compared with mixing ($\sim 2 \times 10^3$ yr) in the open ocean ensures that its $^{34}\text{S}/^{32}\text{S}$ composition is spatially uniform at any one time¹³. The range of $\delta^{34}\text{S}$ values for a specific time period is determined by the isotopic compositions of marine evaporates of that age¹³. In a closed system the sulphate reservoir can become isotopically heavier in S by bacterial reduction. This involves bacterial excretion of H_2S which may produce sulphide which is 10–60% depleted in ^{34}S relative to the original sulphate⁸. In the presence of reduced iron in solution, the H_2S will react to form FeS_2 . Subsequent oxidation of this H_2S (or FeS_2) therefore can produce sulphate containing lighter S than that in the original sulphate reservoir. The crystallization of sulphate evaporites only involves an enrichment of 1–2% in ^{34}S relative to dissolved sulphate¹³. If a similar fractionation is assumed for the crystallization of francolite, much larger differences between its structural $\delta^{34}\text{S}$ composition and seawater values can be explained by the aforementioned processes.

Whilst the sulphate reservoir becomes isotopically heavy in sulphate reducing conditions, the bicarbonate reservoir becomes isotopically light due to inputs of organically derived carbon ($\delta^{13}\text{C}_{\text{PDB}} = -25\%$ approximately)¹⁴. Authigenic francolite precipitated during early diagenesis should therefore contain heavy S and light C. Nearer the sediment–water interface where H_2S is oxidized to give isotopically light porewater sulphate, porewater bicarbonate becomes isotopically heavier in C by mixing with seawater bicarbonate ($\delta^{13}\text{C}_{\text{PDB}} = -0\%$)¹⁵; the heaviest $\delta^{13}\text{C}$ value yet found in authigenic phosphorites (nine samples) is -6.8% PDB (refs 4, 6). The formational environment imposes less influence on the $\delta^{13}\text{C}$ composition of replacement phosphorites due to the masking effect of the precursor carbonate. Sample 1022(1) (Table 1) has the lightest $\delta^{13}\text{C}$ value (-3.8% PDB) yet found in unweathered phosphatized carbonates (40 samples)^{3,4,6}. The $\delta^{34}\text{S}$ composition of replacement phosphorites indicates whether phosphatization occurred in sulphate reducing conditions (heavy $\delta^{34}\text{S}$), open seawater (seawater $\delta^{34}\text{S}$) or at the sulphate reducing/oxic porewater boundary (light $\delta^{34}\text{S}$). The possible combinations of $\delta^{34}\text{S}$ and $\delta^{13}\text{C}$ values in sedimentary francolite are shown in

Fig. 1 with a crude interpretation of formational environment. The $\delta^{13}\text{C}$ values in this simplified model are arbitrary limits only, based on data^{3,4,6} on phosphorites exhibiting other evidence for either an authigenic or replacement origin.

Francolite samples from nine different phosphate deposits were used in this study. Each specimen is similar petrographically, in bulk geochemistry, and its $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ composition to other samples from the same area⁶ and was chosen to typify their deposit. The samples were leached of extraneous carbonate³ and CO_2 prepared for isotopic analysis by reacting with 100% phosphoric acid *in vacuo*¹⁶. For sulphur isotope analysis the samples were left in contact with distilled water for a week to leach soluble sulphates, digested in 10% hydrochloric acid and the sulphur precipitated out as BaSO_4 by adding BaCl_2 solution. SO_2 was prepared by thermal decomposition of BaSO_4 at $1,120^\circ\text{C}$ *in vacuo*¹⁷. Isotopic measurements of the resultant CO_2 and SO_2 gases were made on two separate VG Micromass mass spectrometers, a 903 triple collector and a 602-C double collector, respectively. Samples were run with respect to laboratory standards CP1 for $\delta^{34}\text{S}$ determinations and MCS for $\delta^{13}\text{C}$ analyses. CP1 is a chalcopyrite calibrated with respect to Cañon Diablo Meteorite and MCS is a calcite calibrated with respect to the international carbonate standard PDB and carbon isotope standards TKL 1, K2, NBS 21 and NBS 22 (ref. 18). MCS has also been measured with respect to the new reference materials

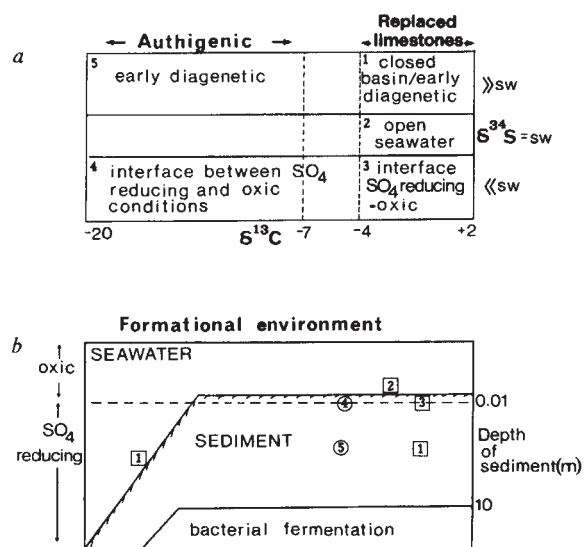


Fig. 1 *a*, Characterization (1–5) of the formational environment of francolite by the isotopic composition of its structural carbonate and sulphate. *b*, Schematic model relating formational environments (1–5) to sedimentary facies. ○, Authigenic; □, replaced carbonates.

NBS 18 and 19 for which there are no accepted values but our preliminary determinations give -5.15% PDB and $+1.87\%$ PDB respectively.

The results presented in Table 1 show that the phosphorites from the Agulhas Bank, Blake Plateau, Californian Borderland and Chatham Rise, which are all considered to be phosphatized carbonates from petrographic⁶ and carbon isotope studies⁴⁻⁶, all have $\delta^{34}\text{S}$ values similar to sulphate evaporites of the same age. This suggests that the phosphatizing medium was open seawater. The phosphatized carbonates from the Moroccan Shelf are slightly enriched in ^{34}S relative to the seawater at the time of formation. This is the result of bacterial sulphate reduction. One of these samples (1022(1)) contains an appreciable amount of light organically derived carbon and is associated with pyrite which is also a by-product of this process. It is impossible to tell whether the upper boundary of the sulphate reduction zone was within the sediment or above as in an anaerobic basin.

The sample from the Phosphoria Formation is also enriched in ^{34}S compared with the supposed seawater value at the time of formation. Our value ($+15.3\%$) compares well with that of previous workers ($+15.0\%$)⁸ who have analysed a suite of samples from this deposit and interpreted their data as evidence for genesis in an early diagenetic reducing environment. This is consistent with the light carbon isotopic composition of these phosphorites^{5,6}.

The francolite samples from the Peru/Chile Shelf, Namibian Shelf (pelletal) and onshore South Africa all contain structural sulphur which is isotopically much lighter than that of the seawater during formation. This implies that oxidation of H_2S occurred before the sulphate was incorporated into the francolite structure. This can only occur at the interface of the sulphate reduction zone and the overlying suboxic or oxic zone but below open access to oxic seawater. This may account for the fact that some of the Peru/Chile phosphorites have been reported to form at the periphery of the zone where the sediments are anoxic¹⁹. The $\delta^{13}\text{C}$ data for the offshore deposits of Peru/Chile and Namibia (pelletal) suggest that they are phosphatized carbonates⁴⁻⁶. Dissolution of these carbonates may have been aided by the local increase in acidity on oxidation of the H_2S . There is a possibility that the onshore South African samples have light sulphur isotopic compositions due to weathering⁸. If this were the case one would expect sample SA/V3(C)/79 to have a lighter $\delta^{34}\text{S}$ value than sample SA/V7(B)/79 which has been less weathered. Since it is the other way round, these $\delta^{34}\text{S}$ values are thought to be essentially unaltered.

The recent concretion from the Namibian Shelf has a slightly lighter $^{34}\text{S}/^{32}\text{S}$ composition than present-day seawater. This may also be the result of oxidation of H_2S which is known to be present in these sediments¹². This sample contains a substantial amount of isotopically light carbon ($\delta^{13}\text{C} = -8.9\%$) derived from the bacterial sulphate reduction process but only a small depletion of ^{34}S relative to seawater. This probably reflects the varying sizes of the seawater reservoirs; the bicarbonate reservoir being small (140 mg l^{-1})¹⁰ and thus isotopically sensitive to small inputs whereas the sulphate reservoir is larger ($2,712 \text{ mg l}^{-1}$)¹⁰ and therefore isotopically less sensitive to inputs.

The sulphur composition of sedimentary francolite has great potential in unravelling past environments of formation, especially when used in conjunction with its structural carbon isotopic composition. This study alone suggests that both authigenic francolite and phosphatized carbonate can form in both sulphate reducing conditions and more oxic conditions. It has also pin-pointed the depth of formation in the sediment column for the francolite samples from Namibia, onshore South Africa and offshore Peru/Chile to the sulphate reducing-oxic porewater interface. It is interesting to speculate that the oxidation of H_2S , referred to above, would produce a slight increase in acidity which may favour precipitation of phosphate relative to carbonate²⁰ and be a pertinent environmental control.

The data and contributions of R.A.B. and M.L.C. are published with the approval of the director, Institute of Geological Sciences (NERC). We thank NERC for the research studentship to R.A.B. and for grant GR3/3819 to J.M.M. which have made this work possible. We also thank Miss S. Liddiard for preparing the typescript. This paper is a contribution to IGCP project 156.

Received 28 September 1982; accepted 20 January 1983.

1. Bremner, J. M. *J. geol. Soc. Lond.* **137**, 773-786 (1980).
2. Cullen, D. J. *SEPM Spec. Publ.* No. 29, 139-148 (1980).
3. Kolodny, Y. & Kaplan, I. R. *J. sedim. Petrol.* **40**, 954-959 (1970).
4. McArthur, J. M., Coleman, M. L. & Bremner, J. M. *J. geol. Soc. Lond.* **137**, 689-673 (1980).
5. Benmore, R. A., McArthur, J. M. & Coleman, M. L. *Proc. Int. Seminar and Field Workshop on phosphorite* (IGCP-156), Udaipur, India (in the press).
6. Benmore, R. A. thesis, Univ. College London (in preparation).
7. Bliskovskiy, V. Z., Grinenko, V. A., Migdisov, A. A. & Savina, L. I. *Geochem. Int.* **8**, 148-155 (1978).
8. Nathan, Y. & Nielsen, H. *SEPM Spec. Publ.* No. 29, 73-78 (1980).
9. Longinelli, A. & Nutti, S. *Earth planet. Sci. Lett.* **5**, 13-16 (1968).
10. Culk, F. in *Chemical Oceanography* (eds Riley, J. P. & Skirrow, G.) (Academic, London, 1965).
11. Redfield, A. C. *Am. Sci.* **46**, 202-226 (1958).
12. Baturin, G. N. *Oceanology*, **12**, 849-855 (1972).
13. Claypool, G. E., Holser, W. T., Kaplan, I. R., Saka, J. & Zak, I. *Chem. Geol.* **28**, 199-260 (1980).
14. Irwin, H., Curtis, C. & Coleman, M. L. *Nature* **269**, 209-213 (1977).
15. Craig, H. *Geochim. cosmochim. Acta* **3**, 53-92 (1953).
16. McCrea, J. M. *J. chem. Phys.* **18**, 849-857 (1950).
17. Coleman, M. L. & Moore, M. P. *Analyt. Chem.* **50**, 1594 (1978).
18. Schoell, M., Faber, E. & Coleman, M. L. *Org. Geochem.* (in the press).
19. Burnett, W. C., Veeh, H. H. & Soutar, A. *SEPM Spec. Publ.* No. 29, 61-71 (1980).
20. Nathan, Y. & Sass, E. *Chem. Geol.* **34**, 103-111 (1981).

Possible limits on the composition of the Archaean ocean

James C. G. Walker

Space Physics Research Laboratory, Department of Atmospheric and Oceanic Science, The University of Michigan, Ann Arbor, Michigan 48109, USA

It has been suggested that the partial pressure of carbon dioxide in the terrestrial atmosphere was larger in the past than it is at present¹⁻⁴. In particular, partial pressures of 100-1,000 times those of the present have been invoked for the early Earth as a means of insuring equable climates at a time when the Sun was significantly less luminous⁵⁻⁷. While the climatic argument is not conclusive⁸, it is worth considering whether the geological record is in any way inconsistent with the proposed high partial pressures of carbon dioxide. I now examine the potential impact of high carbon dioxide partial pressure on ocean chemistry and ask what constraints are imposed by the known record of chemical sedimentation through time. The evidence consists of the persistence throughout almost the entire sedimentary rock record of calcium carbonate and sulphate precipitation. I adopt a uniformitarian point of view that assumes no very great change in the conditions for the deposition of these chemical sediments. The methods of Holland⁹ are used to set limits on the composition of the water from which precipitation occurred. I find no inconsistencies between the sedimentary rock record and presumed higher partial pressure of carbon dioxide early in Earth history, provided that high partial pressure was accompanied by a generally lower pH for seawater, higher concentrations of calcium and bicarbonate ions, and lower concentrations of carbonate and sulphate ions.

For specified values of the pH and carbon dioxide partial pressure the concentrations of carbonate ions and bicarbonate ions in solution are determined by equilibrium relationships among the carbon-bearing species¹⁰. At a temperature of 25 °C and pressure of a few atmospheres the relationships are

$$[\text{H}^+][\text{HCO}_3^-] = 10^{-7.5}P \quad (1)$$

$$[\text{H}^+]^2[\text{CO}_3^{2-}] = 10^{-16.6}P \quad (2)$$

where $[X]$ denotes the concentration of dissolved species X in mol l^{-1} and P denotes the partial pressure of carbon dioxide in

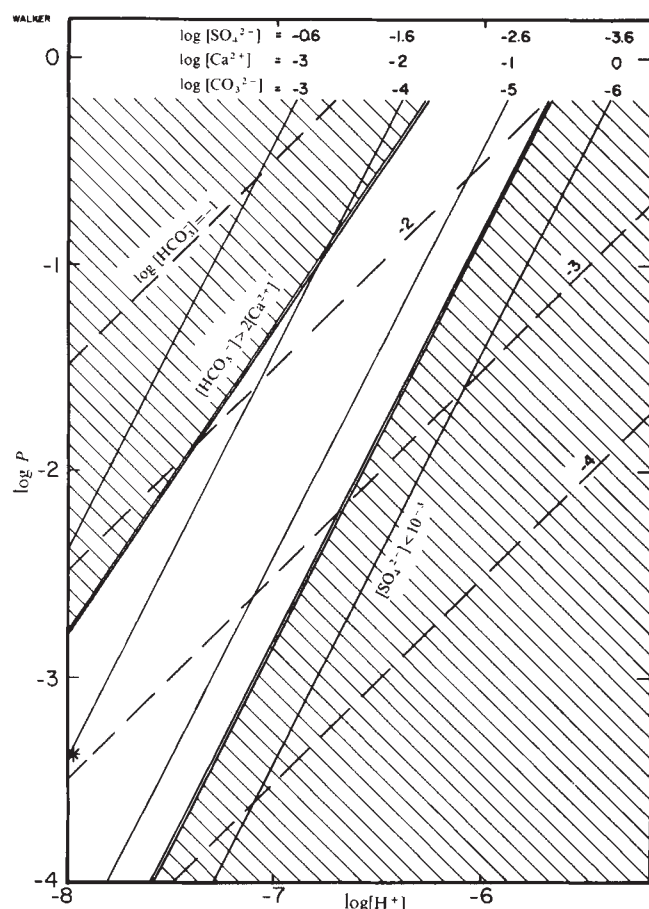


Fig. 1 Concentrations of dissolved species in seawater as functions of the carbon dioxide partial pressure and the concentration of H^+ ions. Concentrations are expressed in mol l^{-1} and the partial pressure in atmospheres. Equilibrium in the carbonate system yields the dashed lines for bicarbonate concentration and the solid lines for carbonate concentration. The values for calcium and sulphate concentrations are derived by assuming a constant product of calcium and carbonate concentrations and of calcium and sulphate concentrations. The star indicates the composition of present day seawater. The cross-hatched regions correspond to compositions that are not consistent with the sedimentary rock record.

atmospheres¹¹. I ignore the weak dependence of the equilibrium constants on pressure and temperature. Constant carbonate and bicarbonate ion concentrations are shown as solid and dashed lines in Fig. 1.

Deposits of calcium carbonate are present in sedimentary rocks of all ages extending back to the oldest yet discovered¹². From the persistence of calcium carbonate precipitation in the sedimentary rock record I estimate that the ocean has always been close to saturation with respect to calcium carbonate, as it is today. With the assumption that the product of calcium and carbonate ion concentration has not changed markedly with time, then,

$$[Ca^{2+}][CO_3^{2-}] = 10^{-6} \quad (3)$$

where values for the present day oceanic average are as quoted by Broecker¹³. In Fig. 1, therefore, the lines of constant carbonate ion concentration apply also to the calcium ion concentration, with the values shown.

Sedimentary sulphate evaporites are known from rocks on all the continents except Antarctica and of all ages extending back to 3,500 Myr ago¹⁴. These sulphate evaporites include massive deposits tens of metres thick and covering thousands of square kilometres in horizontal extent. Many of them contain evidence of having been initially precipitated as gypsum ($GaSO_4 \cdot 2H_2O$). With allowance for the ease with which evaporite deposits may have been destroyed by weathering

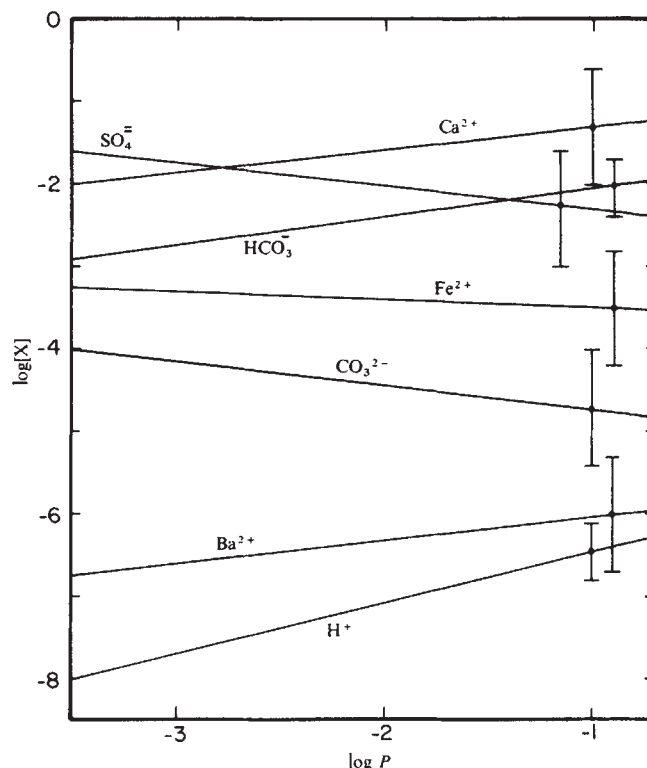


Fig. 2 Seawater composition consistent with the rock record as a function of carbon dioxide partial pressure. The lines correspond to the middle of the allowed region of Fig. 1. The bars represent the range of allowed concentrations. The barium ion concentration corresponds to the assumption of constant product of barium and sulphate concentrations. The iron concentration corresponds to an ocean saturated with respect to siderite.

processes, the record suggests that the mineralogy of the initial stages of evaporite formation has not changed significantly with time. The record is at least consistent with the uniformitarian assumption that gypsum was the first abundant mineral, after carbonate, to precipitate from evaporating seawater in the Precambrian as in the Phanerozoic¹⁵. The record is not sufficiently good to determine whether the proportions of carbonate and gypsum have changed with time. It is therefore conservative to assume that the product of calcium and sulphate ion concentrations has not changed with time. Then

$$[Ca^{2+}][SO_4^{2-}] = 10^{-3.6} \quad (4)$$

and the lines of constant carbonate concentration in Fig. 1 apply also to sulphate, with values as shown. The oxidation state of sulphur in the Archaean ocean and atmosphere has been discussed by Brimblecombe and Walker^{16,17}. Even in the absence of free oxygen it seems likely that photochemical reactions in the atmosphere and surface ocean would have yielded abundant sulphate from volcanic emanations of hydrogen sulphide and sulphur dioxide. The banded iron-formations of the Archaean and early Proterozoic provide evidence for abundant ferrous ions in solution¹⁸⁻²¹, which argues against high sulphide concentrations because of the low solubility of iron sulphides²².

Some constraints on the composition of seawater follow from the occurrence of massive deposits of evaporitic sulphate minerals some tens of metres thick. An approximate lower limit on the sulphate concentration can be derived from the requirement that such a deposit accumulate in a reasonable period of time. A generous upper limit on the evaporation rate in arid terrestrial environments is 10 m yr^{-1} (ref. 23). At this rate and at a sulphate concentration of $10^{-3} \text{ mol l}^{-1}$ it would take 13,500 yr to accumulate a layer of gypsum 1 m thick. The stability of evaporitic environments for times longer than 100,000 yr seems unlikely, so sulphate concentrations $< 10^{-3} \text{ mol l}^{-1}$ are probably not consistent with the geological record. The excluded compositions are indicated by shading on

the right of Fig. 1. A lower limit on carbonate ion concentration and upper limit on calcium ion concentration correspond to the lower limit on sulphate ion concentration.

If evaporating seawater is to precipitate first calcium carbonate and then calcium sulphate then the calcium ion concentration must exceed one half the bicarbonate ion concentration⁹. If this were not the case, precipitation of calcium carbonate would exhaust the calcium ions in seawater leaving none to enter gypsum. The persistence of calcium sulphate evaporites in the sedimentary rock record therefore sets an upper limit on the bicarbonate ion concentration indicated by the shaded area on the left of Fig. 1. The clear band in the middle of Fig. 1 depicts compositions of seawater that are consistent with the sedimentary rock record in the sense that they would have yielded no obvious secular changes in the minerals deposited during the initial stages of evaporite formation. The allowed concentrations are summarized as functions of carbon dioxide partial pressure in Fig. 2. Higher carbon dioxide partial pressure on the early Earth would have corresponded to generally higher concentrations of calcium, bicarbonate and hydrogen ions and generally lower concentrations of carbonate and sulphate ions. Cameron²⁴ has argued from sulphur isotope data that sulphate concentrations were $<10^{-3} \text{ mol l}^{-1}$ in the Archaean ocean.

Modern seawater is close to saturation with respect to the mineral barite (BaSO_4) (ref. 10). The concentrations of barium ions shown in Fig. 2 are derived assuming that the product of barium and sulphate ion concentrations has not changed with time. The generally higher barium concentrations that correspond to higher carbon dioxide partial pressures are consistent with the deduction of Veizer *et al.*²⁵ that the concentration of barium ions was enhanced by about a factor of 10 in the Archaean ocean. The concentrations of iron ions are those that yield saturation with respect to the mineral siderite (FeCO_3), calculated in accordance with the factors described by Holland¹⁹. High concentrations (exceeding $10^{-4} \text{ mol l}^{-1}$) of dissolved ferrous iron in Archaean and early Proterozoic seawater²⁶ are favoured as a source of the chemically precipitated iron in banded iron formations¹⁸⁻²¹.

In the absence of silica-precipitating organisms, the oceans were presumably approximately saturated with respect to amorphous silica. Holland⁹ estimates a concentration of dissolved silica of about 100 p.p.m. This high concentration of dissolved silica is presumably responsible for the abundance of chemically precipitated siliceous sediments in formations of Archaean and early Proterozoic age. In the absence of extensive continental land masses²⁷ there were presumably few large deposits of evaporitic halite (NaCl). Enhanced concentrations of sodium and chloride ions in Archaean seawater are a possibility. The oceans would be far from saturation with respect to halite even if all known evaporite deposits were dissolved¹⁵.

I have not discussed possible reasons for the postulated high pressures of carbon dioxide on the early Earth. Mechanisms that relate to the rate of weathering of continental rocks^{6,28,29} cannot be invoked for the Archaean era, when oceanic composition was dominated by interactions with the mantle rather than with continental rocks²⁵. In general terms, the carbon dioxide partial pressure would have been determined by the partitioning of carbon between the fluid (ocean plus atmosphere) and solid (rock) phases. Higher internal temperatures and more tectonic activity on the early Earth may have promoted the remobilization of carbon from the solid to the fluid phases^{1,30}. In the conditions described here, the fraction of total terrestrial carbon in the fluid phase is still small. A more precise understanding of the controls on the composition of the Archaean ocean and atmosphere undoubtedly involves the details of the interaction of seawater with mantle rocks^{25,27,31}.

This paper describes research supported in part by NASA under grant NAGW-176 to the University of Michigan. It is a contribution of the Precambrian Paleobiology Research Group.

Received 26 November 1982; accepted 20 January 1983.

1. Holland, H. D. in *The Early History of the Earth* (ed. Windley, B. F.) 559-567 (Wiley, London, 1976).

2. Hart, M. H. *Icarus* **33**, 23-39 (1978).
3. Henderson-Sellers, A. & Cogley, J. G. *Nature* **298**, 832-835 (1982).
4. Veizer, J. *Contr. Miner. Petrol.* **38**, 261-278 (1973).
5. Owen, T., Cess, R. D. & Ramanathan, V. *Nature* **277**, 640-642 (1979).
6. Walker, J. C. G., Hays, P. B. & Kasting, J. F. *J. geophys. Res.* **86**, 9776-9782 (1981).
7. Walker, J. C. G. *Paleogeogr. Paleoclimatol. Paleocool.* **40**, 1-11 (1982).
8. Rossow, W. B., Henderson-Sellers, A. & Weinreich, S. K. *Science* **217**, 1245-1247 (1982).
9. Holland, H. D. *Geochim. cosmochim. Acta* **36**, 637-651 (1972).
10. Riley, J. P. & Chester, R. *Introduction to Marine Chemistry* (Academic, New York, 1971).
11. Walker, J. C. G. *Evolution of the Atmosphere* (Macmillan, New York, 1977).
12. Allaart, J. H. in *The Early History of the Earth* (ed. Windley, B. F.) 177-189 (Wiley, London, 1976).
13. Broecker, W. S. *Chemical Oceanography* (Harcourt Brace Jovanovich, New York, 1974).
14. Walker, J. C. G. *et al.* in *The Earth's Earliest Biosphere: Its Origin and Evolution* (ed. Schopf, J. W.) (Princeton University Press, in the press).
15. Holland, H. D. *The Chemistry of the Atmosphere and Oceans* (Wiley-Interscience, New York, 1978).
16. Brimblecombe, P. & Walker, J. C. G. *J. geophys. Res.* (submitted).
17. Brimblecombe, P. & Walker, J. C. G. *J. geophys. Res.* (submitted 1982).
18. Drever, J. I. *Bull. geol. Soc. Am.* **85**, 1099-1106 (1974).
19. Holland, H. D. *Econ. Geol.* **68**, 1169-1172 (1973).
20. Ewers, W. E. in *4th int. Symp. Environmental Biogeochemistry* (eds Trudinger, P. A. & Walter, M. R.) (Springer, New York, 1980).
21. Button, A. *et al.* in *Mineral Deposits and the Evolution of the Biosphere* (eds Holland, H. D. & Schidlowski, M.) 259-273 (Springer, Berlin, 1982).
22. Walker, J. C. G. *Life Sci. Space Res.* **8**, 89-100 (1980).
23. Sellers, W. D. *Physical Climatology* (University of Chicago Press, 1965).
24. Veizer, J., Compston, W., Hoefs, J. & Nielsen, H. *Naturwissenschaften* **69**, 173-180 (1982).
25. Cameron, E. M. *Nature* **296**, 145-148 (1982).
26. Veizer, J. in *The Earth's Earliest Biosphere: Its Origin and Evolution* (ed. Schopf, J. W.) (Princeton University Press, in the press).
27. Lovelock, J. E. & Whitfield, M. *Nature* **296**, 561-563 (1982).
28. Lovelock, J. E. & Watson, A. J. *Planet. Space Sci.* **30**, 795-802 (1982).
29. Walker, J. C. G. in *Comparative Planetology* (ed. Ponnamperna, C.) 141-163 (Academic, New York, 1978).
31. Fryer, B. J., Fyfe, W. S. & Kerrich, R. *Chem. Geol.* **24**, 25-33 (1979).

Seasonal sedimentation of phytoplankton to the deep-sea benthos

D. S. M. Billett, R. S. Lampitt & A. L. Rice

Institute of Oceanographic Sciences, Wormley, Godalming, Surrey GU8 5UB, UK

R. F. C. Mantoura

Institute for Marine Environmental Research, Prospect Place, The Hoe, Plymouth, Devon PL1 3DH, UK

Until recently the deep sea was considered to be a particularly stable environment¹, free from seasonal variations. However, atmospheric storms may cause periodicity in deep-ocean currents² and nepheloid layers³ while seasonality in the particulate flux to the deep sea is known to occur in the Sargasso Sea^{4,5} and Panama Basin⁶. Evidence is presented here of a similar seasonal pulse of detrital material to bathyal and abyssal depths in temperate latitudes; this material seems to be derived directly from the surface primary production and to sink rapidly to the deep-sea benthos. Considerable sedimentation occurs soon after the spring bloom and continues throughout the early summer. This process acts as a pathway for the descent of carbon from the euphotic zone, providing a periodic food source for the deep pelagic and benthic communities.

Since 1979, routine photo-transects obtained with the Institute of Oceanographic Sciences' epibenthic sledge⁷⁻⁹, and time-lapse photography using the Bathysnap system¹⁰, have revealed a patchy, detrital layer on the sea floor in the Porcupine Seabight in the north-east Atlantic (50° N, 13° W), at depths between 1,370 and 4,100 m, during the 4 months from April to July. Six cruises conducted at other times of the year have failed to detect the presence of the material. On 1 April 1982, Bathysnap was deployed at 2,000 m to take hourly photographs over a 33-day period. Detrital particles began to accumulate on the sediment surface from 24 April, but there was no appreciable sedimentation until 30 April, when the detrital layer increased rapidly over a 38-h period (Fig. 1). No substantial increase occurred during the remaining 2 days of observation, but considerably more of the sediment surface was covered in the same locality by the end of May in both 1981 and 1982. A further deployment at the end of July 1982 for 30 days at

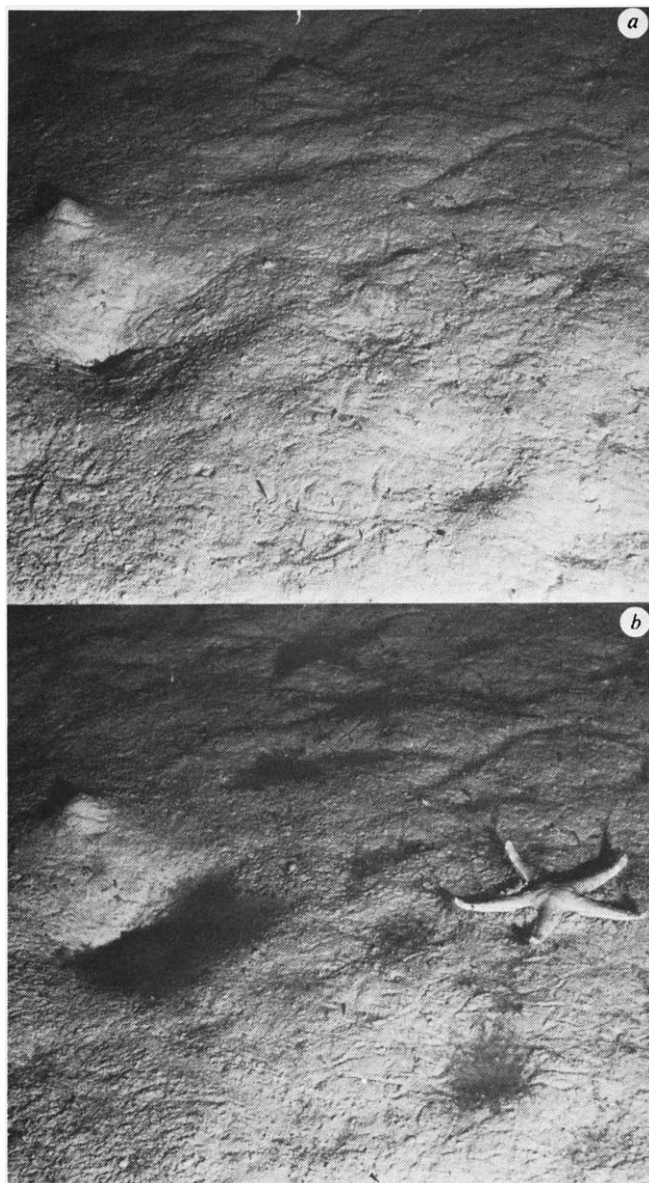


Fig. 1 Bathysnap photographs of the same area of sea bed taken; *a*, at 15.20 h on 29 April 1982 and *b*, 38 h later. The area is 0.6 m² and the arm radius (*R*) of the asteroid, *Plutonaster bifrons*, is 80 mm. The detritus accumulates in depressions on the sediment surface and is seen in these photographs as dark patches.

1,300 m showed a steady decline in the quantity of detritus on the seabed.

All the nearbed currents (1.3 m above) recorded by Bathysnap in the Porcupine Seabight have been tidal with speeds up to 13 cm s⁻¹. Although similar at all depths, the effect of these currents on the detrital layer observed in the photographs has been very variable. At some stations there was considerable movement of the material with substantial periodical loss of visibility. At another station, however, individual aggregates (2–5 mm in diameter) which arrived on the seabed did not move during the period of observation (<2 days).

Undisturbed samples of the sediment surface (Fig. 2) were obtained using the Scottish Marine Biological Association's Multiple Corer¹¹. A sample taken in May 1981 at 2,000 m consisted principally of diatoms, in some instances bound with amorphous organic material into small aggregations. The species assemblage was typical of spring oceanic phytoplankton (A. W. G. John, Institute for Marine Environmental Research, personal communication). Intensive coring in July 1982 obtained more detrital material at depths from 1,340 to 4,475 m. Diatomaceous material was again present but faecal

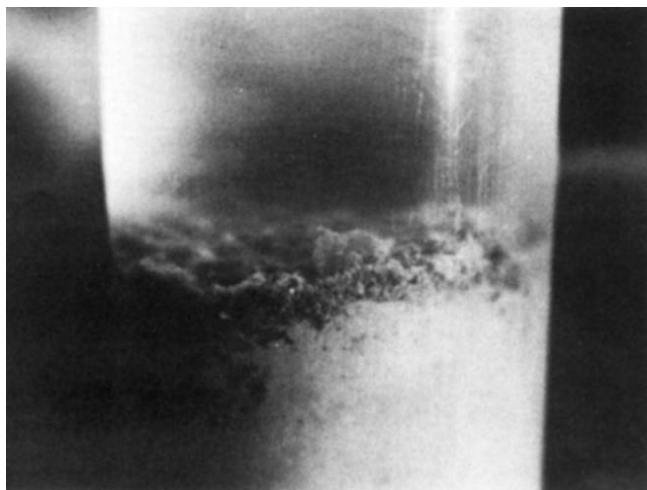


Fig. 2 A sub-core (58 mm diameter) from the SMBA Multiple Corer designed by Dr P. R. O. Barnett and Mr J. Watson. The detrital layer can be seen as a dark band, about 5 mm thick, on top of the sediment. The water overlying the sediment is clear of resuspended material, demonstrating the ability of the gear to obtain virtually undisturbed cores.

pellets and discrete gelatinous aggregations up to 10 mm in diameter also occurred; the latter contained many coccolithophorids as well as dinoflagellates, crustacean eggs, small faecal pellets and amorphous organic material. In this case the species composition was typical of summer phytoplankton (A. W. G. John, personal communication) indicating that significant sedimentation occurs in early summer as well as in spring. A similar temporal change in the constituents of the particle flux was noted in the Panama Basin⁶.

Detrital material sampled in July 1982 was analysed for organic carbon and nitrogen using a low temperature oxygen plasma asher for the selective oxidation of organic matter, and a Carlo Erba 1106 CN analyser. Chlorophyll and carotenoid pigments from the same material were extracted in 90% acetone and analysed by ion-paired reverse-phase HPLC¹². Qualitatively, the HPLC absorbance and fluorescence chromatograms were similar for all six samples taken between 1,362 and 4,475 m (Fig. 3), although the total chloropigment content of the detritus decreased by two orders of magnitude with increasing depth (R.F.C.M., unpublished data). Chlorophyll *a* accounted for only 5.4–17.8% (mean 10.5%) of the total chloropigments. The virtual absence of chlorophyll *b* indicates there was little, if any, terrestrial input, whereas the presence of both chlorophyll *c* and fucoxanthin-type carotenoids¹³, determined from the absorbance chromatograms, confirms visual observations that the detrital material is derived principally from the Bacillariophyceae (diatoms) and the Dinophyceae (dinoflagellates). Phaeophorbide *a* was the most abundant chloropigment, suggesting that most of the detritus or its precursors had been ingested by pelagic or benthic fauna. However, two unidentified compounds in the fluorescence chromatograms together accounted for 22–40% of the total chloropigments, and these may have been formed by bacterial degradation.

The organic carbon and nitrogen content of the detritus is very low (*C* = 5.98% dry weight ± 0.66 s.e., *N* = 0.25% dry weight ± 0.04 s.e.) compared with values for living phytoplankton^{14,15}. Although this indicates extensive degradation the values are still substantially higher than for deep-sea superficial sediments^{16,17}. The *C*:*N* weight ratio of 24 suggests preferential degradation of nitrogenous compounds and, although higher than values for some sediment trap material^{18,19}, is not without precedent^{20,21}.

The close similarity between the phytoplankton species assemblages of the detritus and surface water before the date of sampling indicates rapid transport of detrital material. Since strong nearbed currents have not been detected in the

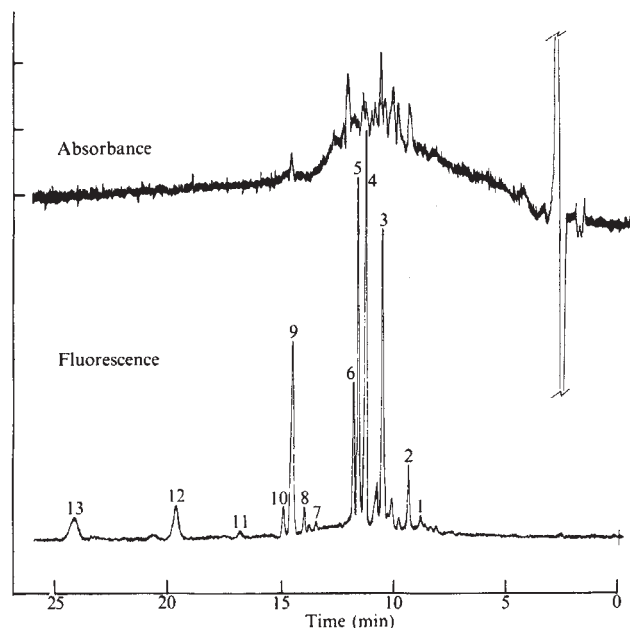


Fig. 3 HPLC absorbance (440 nm; 1 div. = 0.001 AU) and fluorescence (λ_{ex} = 630 nm, λ_{emm} > 600 nm) chromatograms of 90% acetone extracts of benthic detritus from St. 51612, 2,670 m, 50°19' N, 13°21' W. The identities of the fluorescent peaks are: 1, chlorophyllide *a*; 2, chlorophyll *c*; 3, unknown type *a* chloropigment; 4, phaeophorbide *a*; 5, unknown type *a* chloropigment; 6, phaeophorbide *c*; 7, chlorophyll *b*; 8, bacteriochlorophyll (?); 9, chlorophyll *a*; 10, chlorophyll *a'*; 11, phaeophytin *b*; 12, phaeophytin *a*; 13, pyrochlorophyll *a*.

Porcupine Seabight, vertical rather than horizontal transport processes are indicated. Assuming that the spring bloom in the Porcupine Seabight occurred in early April 1982 (Continuous Plankton Recorder data, Institute for Marine Environmental Research), our results suggest that diatoms can sink to 2,000 m in as little as 2–3 weeks (approximately 100–150 m day⁻¹). This is considerably faster than would be expected from laboratory experiments on monospecific algal cultures²². In shallow water, unexpectedly high sinking rates have also been reported²³ and a number of mechanisms for this have been proposed²². One of these, the incorporation of diatoms into zooplankton faeces, may be important in the summer months but is unlikely to be significant at the start of the spring bloom when zooplankton biomass in temperate waters is often low. This is endorsed by the presence of intact diatom chains and the scarcity of faecal pellets in the detritus collected in May 1981. Alternatively, the temporary accumulation of diatoms at density discontinuities may lead to the formation of faster sinking aggregates. This process has been proposed to account for the high sinking rates of *Skeletonema* in shallow water²⁴, and is supported by diver observations in the Sargasso Sea where detrital material collecting at the thermocline has been seen to peel off and sink in strings up to 15 cm long (V. McAlister, personal communication). In contrast, the gelatinous aggregations could be formed by zooplankton, and may be analogous to the fast-sinking coccolithophorid-associated mucus aggregations observed in the Panama Basin⁶.

The seasonal sedimentation of phytoplankton and its effect on the benthos have been studied in shallow temperate seas^{25,26}. Deep-sea organisms may respond to the arrival of the detritus in a similar fashion to the shallow-water fauna. Synchronous annual reproductive cycles with spawning in the early spring have been found in several deep-sea echinoderms which produce small eggs, indicative of indirect planktonic development²⁷. The food source for the seasonal production of planktotrophic, and possibly demersal larvae, could be produced by the sedimentation of diatoms from the spring bloom. Behavioural responses to the benthic detritus by megafaunal invertebrates have been noted. The echinoid *Echinus affinis*

appears to forage actively for detrital accumulations (R.S.L., unpublished Bathysnap data), while the incorporation of detrital patches into the sediment may cause the aggregation of the holothurian *Kolga hyalina*²⁸.

Previous results have indicated a marked seasonality in the flux of particulate organic matter from the euphotic zone to the deep sea in both tropical eutrophic⁶ and sub-tropical oligotrophic regions^{4,5}. The existence of a similar seasonal flux in temperate eutrophic waters, reported here, suggests that the phenomenon is widespread and will be of great significance to both bathyal and abyssal benthic communities.

We thank Dr M. V. Angel for comments on the manuscript, Drs M. J. R. Fasham, P. R. Pugh and W. R. Simpson for discussion, Dr P. R. O. Barnett and Mr J. Watson for providing and operating the SMBA Multiple Corer, Mr A. W. G. John for phytoplankton identifications, Ms C. A. Llewellyn for her technical help, and the Master and crew of RRS *Challenger* for their help during sampling.

Received 16 November 1982; accepted 4 February 1983.

- Sanders, H. L. *Am. Nat.* **102**, 243–282 (1968).
- Dickson, R. R., Gould, W. J., Garbutt, P. A. & Killworth, P. D. *Nature* **295**, 193–198 (1982).
- Gardner, W. D. & Sullivan, L. G. *Science* **213**, 329–331 (1981).
- Deuser, W. G. & Ross, E. H. *Nature* **283**, 364–365 (1980).
- Deuser, W. G., Ross, E. H. & Anderson, R. F. *Deep-Sea Res.* **28**, 495–505 (1981).
- Honjo, S. *Science* **218**, 883–884 (1982).
- Aldred, R. G., Thurston, M. H., Rice, A. L. & Morley, D. R. *Deep-Sea Res.* **23**, 167–174 (1976).
- Rice, A. L., Aldred, R. G., Billett, D. S. M. & Thurston, M. H. *Ambio Spec. Rep.* **6**, 59–72 (1979).
- Rice, A. L., Aldred, R. G., Darlington, E. & Wild, R. A. *Oceanologica Acta* **5**, 63–72 (1982).
- Lampitt, R. S. & Burnham, M. P. *Deep-Sea Res.* (in the press).
- McIntyre, A. D. & Warwick, R. M. in *Methods for the Study of Marine Benthos* 2nd edn (eds Holme, N. A. & McIntyre, A. D.) Ch. 7 (Blackwell, Oxford, in the press).
- Mantoura, R. F. C. & Llewellyn, C. *Analytica chim. acta* (submitted).
- Liaaen-Jensen, S. in *Marine Natural Products Chemistry* (eds Faulkner, D. J. & Fenical, W. H.) 239–254 (Plenum, New York, 1977).
- Parsons, T. R., Takahashi, M. & Hargrave, B. *Biological Oceanographic Processes* 2nd edn (Pergamon, New York, 1977).
- Lisitsyn, A. P., Belyayev, Yu. I., Bogdanov, Yu. A. & Bogoyavlenskiy, A. N. *Int. Geol. Rev.* **9**, 253–274 (1967).
- Smith, K. L. *Mar. Biol.* **47**, 337–347 (1978).
- Khripounoff, A. & Sibuet, M. *Mar. Biol.* **60**, 17–26 (1980).
- Rowe, G. T. & Gardner, W. D. *J. mar. Res.* **37**, 581–600 (1979).
- Honjo, S. *J. mar. Res.* **38**, 53–97 (1980).
- Hinga, K. R., Sieburth, J. M. & Heath, G. R. *J. mar. Res.* **37**, 557–579 (1979).
- Knauer, G. A., Martin, J. H. & Bruland, K. W. *Deep-Sea Res.* **26**, 97–108 (1979).
- Smayda, T. J. *Oceanogr. mar. Biol.* **8**, 353–414 (1970).
- Bodungen, B. von, Bröckel, K. von, Smetacek, V. & Zeitzschel, B. *Kleiner Meeresforsch.* **5**, 49–60 (1981).
- Smetacek, V. in *Flux of Energy and Materials in Marine Ecosystems: Theory and Practice* (ed. Fasham, M. J. R.) (Plenum, New York, in the press).
- Smetacek, V. *Ophelia Suppl.* **1**, 65–76 (1980).
- Graf, G., Bengtsson, W., Diesner, U., Schulz, R. & Theede, H. *Mar. Biol.* **67**, 201–208 (1982).
- Tyler, P. A., Grant, A., Pain, S. L. & Gage, J. D. *Nature* **300**, 747–750 (1982).
- Billett, D. S. M. & Hansen, B. *Deep-Sea Res.* **29**, 799–818 (1982).

A new method of scanning electron microscopy for imaging biological tissues

A. Boyde & S. A. Reid

University College London, Department of Anatomy and Embryology, Gower Street, London WC1E 6BT, UK

The scanning electron microscope (SEM) has proved of little value in the examination of material prepared for light microscopic histology. One of the chief reasons for this is that the secondary electron signal used for image formation in routine scanning microscopy derives from the surface of the specimen. In the case of histological material this surface is one which has been severely distorted by processing and cutting procedures. Light microscopy sections can be usefully studied in the SEM if the signal used to form the image derives from a considerable portion of the thickness of the section. Thus the backscattered electron (BSE) image has been successfully used in studying the distribution of dense material or densely staining components several micrometres deep to the surface of dried

sections¹. Such sections are, however, usually mounted on low density (poorly BSE reflecting) non-transparent substrates such as beryllium or carbon, so that matching light microscopy of the same samples is not possible. We report here a method by which histological sections mounted on glass slides can be imaged in the SEM at a resolution higher than that obtained using conventional light microscopy. The method exploits the facts that the ordinary, cheap light microscope slide is strongly cathodoluminescent, yet the standard histological (7 μm) section is of such a mass thickness that it absorbs a significant proportion of electrons with energies (5–20 keV) usually used in biological SEM. Thus the measure of the glass cathodoluminescence signal is the measure of the electron flux passing through the specimen.

If ordinary histological material mounted on glass microscopic slides is used as the starting point, the embedding wax or the balsam or other mounting medium is dissolved in the appropriate solvent (such as xylene or toluene) and the slide washed with $\text{C}_2\text{Cl}_2\text{F}_3$ and dried by solvent evaporation. Alternatively, critical point drying can be utilized. For cells grown on glass slides or coverslips, conventional fixation and freeze drying or critical point drying techniques for SEM are used. The specimen is given a surface conductive coat of Al,

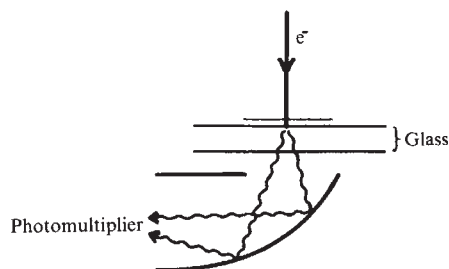


Fig. 1 Diagram showing the arrangement used to produce Fig. 2. The electron beam in an SEM strikes an Al coated, deparaffinized section on a glass slide, generating cathodoluminescence in the glass. This light is directed to the photomultiplier tube normally used for the scintillator detector in the SEM.



Fig. 2 Glass cathodoluminescence SEM image of horse parotid gland, fixed in formol-Zenker, embedded in wax, deparaffinized with xylene, washed with $\text{C}_2\text{Cl}_2\text{F}_3$, air-dried and coated with Al by evaporation. 20kV, 10pA, 40 s. Fieldwidth, 70 μm .

C, Au or Au/Pd. The surface conductive coating may also act as a reflective coating. The cathodoluminescent signal is collected by either a Perspex light guide placed closely beneath the slide; a mirror or prism followed by a lens to focus the light on the photomultiplier window or an Al foil tube which acts as a focusing reflector and light guide (Fig. 1).

We have obtained images from sections of a variety of tissues and of cultured cells grown on glass. These images are obtained with beam currents of ~ 10 pA, only slightly larger than those used for normal SEM imaging and the same as used for BSE imaging on the same material (Fig. 2). Improved contrast is obtained if heavy metals can be included in the fixation recipe, as for example, the Hg and Cr in Zenker's (Fig. 2); but images which are quite satisfactory can also be obtained from formaldehyde-only fixed tissue. The glass cathodoluminescence SEM mode is a kind of scanning transmission electron microscope with no energy filtering. The image shows variations in dry mass distribution and can be compared with BSE imaging, microradiography and interference light microscopy.

The method can be quickly set up and greatly extends the classes of biological material which can be examined in any SEM. It allows the study of glassmounted sections and cells in three dimensions, since stereo pair images can be made by tilting the specimen in the SEM. It will also permit the study of archival histopathological material.

S.A.R. is supported by a research studentship from the Anatomical Society of Great Britain and Ireland.

Received 26 January; accepted 1 March 1983.

1. Becker, R. P. & Geoffroy, J. S. *Scanning Electron Microscopy 1981 Pt IV*, 195–206 (SEM Inc., Chicago, 1981).

Detection of HSV-1 genome in central nervous system of latently infected mice

Daniel L. Rock & Nigel W. Fraser

The Wistar Institute, 36th and Spruce Streets, Philadelphia, Pennsylvania 19104, USA

Herpes simplex virus type 1 (HSV-1) can establish latent infections in peripheral nerve ganglia and the central nervous system (CNS) of experimentally infected mice. Latent infections of peripheral nervous tissue are characterized by the ability to recover infectious virus from explant cultures of most of the latently infected ganglia^{1–4}. In contrast, infectious virus is infrequently recovered from the CNS of latently infected mice following explant culture^{2–4}, although viral DNA can be detected in CNS tissue⁴. The presence of incomplete or defective viral genomes during the latent CNS infections has been proposed to explain this difference between the two latent infections⁴. To characterize the completeness and the physical state of the latent viral genome, we used Southern blot hybridization to analyse the viral DNA from latently infected mouse ganglia and brains. Our results, reported here, indicate that most, if not all, of the HSV-1 genome is present in latently infected CNS tissue and that the latent viral genome exists in a form other than linear, unit length DNA.

Six- to ten-week-old female BALB/c mice were infected in both eyes following moderate corneal scarification with 1.0×10^6 plaque-forming units of HSV-1 strain F (obtained from B. Roizman, University of Chicago). Mortality during the acute phase of the infection was 30–50%, with infectious virus present throughout the CNS. At 6 days post-infection (p.i.), virus was detected using explant culture procedures⁵ in five of five trigeminal ganglia, five of five brain stems (pons-medulla), four of five cerebella, and five of five cerebra. At 2 months p.i., virus could be detected in explant cultures of trigeminal ganglia from 10 of 10 animals, but could not be isolated from the brain stem, cerebellum or cerebrum of these animals. These results are similar to observations reported by others^{3–5}.

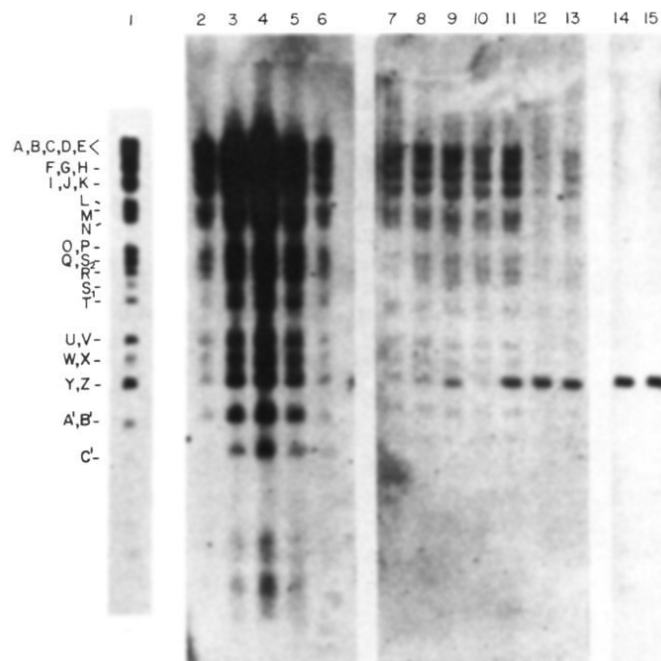


Fig. 1 Detection of HSV-1 DNA sequences in whole mouse brain during the acute and latent phase of infection. DNA extracted from brain tissue was digested with *Bam*HI, electrophoresed in agarose gels, transferred to nitrocellulose filters and hybridized with $\sim 10^8$ c.p.m. of 32 P-labelled HSV-1 strain F DNA. Autoradiographic exposures ranged over 1–6 days. Lane 1, 1.0 ng of HSV-1 DNA, purified from extracellular virions, and 10 μ g of uninfected BHK cell DNA (fragments labelled according to Locker and Frenkel⁷); lanes 2–6, 20 μ g of brain DNA taken from each of five acutely infected mice (6 days p.i.); lanes 7, 8, 10–13, 20 μ g of brain DNA taken from each of six latently infected mice (2 months p.i.); lane 9, 20 μ g of DNA extracted from pooled latently infected trigeminal ganglia (2 months p.i.); lanes 14, 15, 20 μ g of brain DNA taken from each of two uninfected control animals.

DNA was extracted from ganglia and brains of acutely and latently infected mice, digested with restriction endonuclease *Bam*HI, electrophoresed in 0.6% agarose gels, transferred to nitrocellulose filters and hybridized with nick-translated 32 P-labelled viral DNA as described previously⁶. Figure 1 shows the DNA extracted from brains and ganglia of mice during the acute and latent phase of the infection; 1 ng of HSV-1 strain F DNA in 10 μ g of uninfected baby hamster kidney (BHK) cell DNA is shown in Fig. 1, lane 1. Despite the large amount of non-viral DNA in the sample, the pattern of bands is still recognizable as that of HSV-1 strain F⁷. Samples taken from acutely infected CNS tissue (6 days p.i.) show bands corresponding to sequences from all regions of the viral genome (Fig. 1, lanes 2–6). Patterns observed in CNS tissue taken from latently infected animals (2 months p.i.) are shown in Fig. 1, lanes 7, 8 and 10–13. At this level of resolution, the HSV-1 genome patterns observed in acutely infected mouse brain, pooled latently infected trigeminal ganglia and latently infected CNS tissue are indistinguishable (Fig. 1, lanes 2–6, lane 9, lanes 7, 8, 10–13, respectively). Further attempts to determine whether all regions of the viral genome were represented in latently infected brains were made using cloned fragments of the viral genome as probes in blot hybridizations (Fig. 2). All latently infected brains examined were positive for HSV-1 *Bam*HI fragments B, E, N, Q and the *Eco*RI fragment H. In all cases, the fragments detected were of virion size (data not shown).

The defective genomes of HSV-1 observed in tissue culture-infected cells are composed of multiple reiterations of limited regions of the viral genome^{8–10}. The results reported here from blot hybridizations, using total HSV-1 DNA and cloned fragments of the viral genome as probes, demonstrate that most if not all of the viral genome is present in latently infected mouse

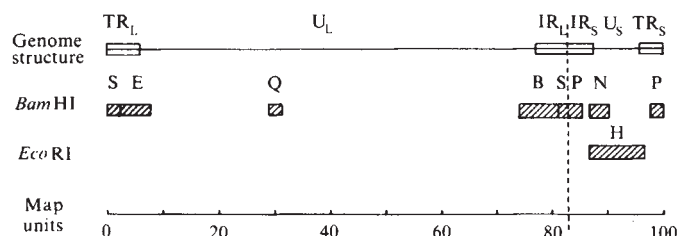


Fig. 2 Cloned viral fragments used as probes in blot hybridization experiments are localized along the genome map of the prototype HSV-1 F genome¹⁷. The genome consists of a long unique region (U_L) and short unique regions (U_S). Both unique regions are bounded by inverted repeat regions (TR_L , IR_L , IR_S , TR_S). The long and short segments of the genome invert relative to one another, forming four isomers. The fragment *Bam*HI SP, which spans the junction of the long and short segment of the genome, is totally within the IR_L and IR_S regions, resulting in a single *Bam*HI junction fragment for all isomers. The genome is divided into 100 map units, each of which represents 0.98×10^6 daltons or ~ 1.63 kb.

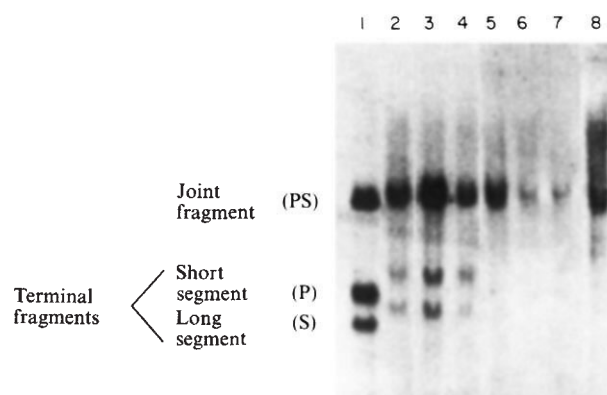


Fig. 3 Blot hybridization of nick-translated 32 P-labelled *Bam*HI SP DNA to *Bam*HI digests of brain DNA taken from acutely and latently infected mice. Lane 1, 1.0 ng of HSV-1 strain F DNA and 10 μ g of uninfected BHK cell DNA; lanes 2–4, 20 μ g of brain DNA taken from each of three acutely infected mice (6 days p.i.); lanes 5–7, 20 μ g of brain DNA taken from each of three latently infected mice (2 months p.i.); lane 8, 20 μ g of DNA extracted from pooled latently infected trigeminal ganglia (5 months p.i.).

brains, and suggest that the inability to reactivate virus from the CNS following explant culture is not due to incomplete or 'defective' representation of the viral genome in the CNS.

Hybridization of 32 P-labelled HSV-1 virion DNA (grown in CV-1 cells) to uninfected mouse brain DNA resulted in a 2 kilobase-pair (kbp) band (Fig. 1, lanes 14, 15), which also appeared in most of the other samples examined (Fig. 1, lanes 7–13), although with varying intensity. The band failed to hybridize with 32 P-labelled CV-1 cell DNA (data not shown). This cross-hybridizing band may represent a cellular sequence with homology to regions of the HSV-1 genome. Homology between portions of the HSV inverted repeat sequences and repetitive sequences in mammalian cell DNA has been demonstrated¹¹. In addition, Maitland *et al.*¹² have shown limited homology between regions of the HSV-2 genome and RNA from human placenta.

Using blot hybridization, HSV-1 DNA was detected in $\sim 50\%$ of the whole mouse brains examined (Table 1) and at a concentration comparable with the results of Cabera *et al.*⁴. However, HSV-1 DNA was detected in the brain stem (pons–medulla) of all the animals (Table 1) when defined regions of the brain were examined individually. The brain stem, which constitutes about 1/15th (50 μ g) of the total brain DNA, seems to have a higher concentration of viral sequences than other brain regions. Dilution of the brain-stem region with other brain regions containing less or no viral DNA probably explains the inability to detect viral DNA in some whole brain samples. Cabera *et al.*⁴ detected the HSV-1 genome in individual brain

Table 1 Detection of the HSV-1 genome in mouse brain by blot hybridization

Time post-infection	Whole brain		Genome equivalents per cell
	Positive samples	Total no. of samples	
6 days	5/5		1.5–0.15
2 months	7/15		0.15–0.015
5 months	8/15		0.15–0.015

Brain regions			
Mouse no.	Brain stem (pons–medulla)		
	Cerebellum	Cerebrum	
068	+	–	+
069	+	–	+
070	+	–	–
071	++	–	–
072	++	–	–

DNA was extracted from whole brain at various times p.i. or from brain regions at 2 months p.i. Samples (20 µg) of brain DNA were digested with *Bam*HI, electrophoresed in agarose gels, transferred to nitrocellulose filters and hybridized with ³²P-labelled HSV-1 strain F DNA. Genome equivalents were approximated by visual comparison of the sample autoradiographs with that of reconstructed standards. + Denotes a sample positive for HSV-1 DNA with 0.15–0.015 genome equivalents per cell; ++ indicates that the sample contained more than 0.15 but less than 1.5 genome equivalents per cell; – denotes samples that were negative for HSV-1 DNA in the assay conditions.

regions from only 30% of mice examined 2 months after corneal inoculation with HSV-1 strain F. The severity of the acute viral infection (30–50% mortality) in our experiments, compared with the 2% mortality obtained by Cabera *et al.*⁴, may account for the higher incidence of brains containing HSV-1 DNA at 2–5 months p.i.

Attempts to detect the terminal fragments of the HSV-1 genome in brain DNA taken from acutely and latently infected animals were made using the *Bam*HI SP fragment (PRB 115, ref. 13) of the viral genome as a probe (Fig. 3). Comparison between the pattern observed with virion DNA (Fig. 3, lane 1) and that obtained with acutely infected brain DNA (Fig. 3, lanes 2–4) shows that the terminal fragments of the viral genome are increased in size by ~300–500 bp and are submolar in concentration in the acutely infected brains. This approximates to the size of an 'a' sequence in the terminal repeat and it is possible that reiteration of the 'a' sequence has occurred, as has been noted in lytic tissue culture infections¹³. Densitometer scans of the virion and acute brain autoradiographs showed a molar ratio relative to the PS joint fragment of approximately 0.3 for the P and S terminal fragments during the acute infection. This is not due to isomerization of the viral genome because *Bam*HI cuts within the terminal repeat and, therefore, each isomer gives the same terminal fragment (see Fig. 2). Reduction in the molarity of the virus terminal fragments has been observed in HSV-1-infected tissue culture cells and is thought to be due to the generation of concatemers during viral DNA replication¹⁴.

Latently infected brain DNA (Fig. 3, lanes 5–7) and DNA extracted from pooled latently infected trigeminal ganglia (Fig. 3, lane 8) seem to lack the terminal fragments of the viral genome. It is unlikely that this is due to lack of sensitivity because a latently infected brain (Fig. 3, lane 5) and acutely infected brains (Fig. 3, lanes 2, 4) show PS joint fragments of similar intensity while differing in the presence of the terminal fragments. The absence of terminal fragments of the HSV-1 genome in latently infected tissue argues against any model where the latent viral genome would persist as unit length linear DNA. Rather, these results accommodate any of several other models: (1) persistence of the viral genome in latently infected cells as a closed circular structure, similar to those seen in cells persistently infected with Epstein–Barr virus¹⁵; (2) integration of the viral DNA into cellular DNA via a bacteriophage λ-like mechanism¹⁶ or a random integration mechanism; (3) per-

sistence of the latent viral DNA in some type of concatenated structure such as the high molecular weight replicative intermediates seen during HSV DNA synthesis¹⁴. The results presented here clearly show that the viral DNA inside the latently infected cell is modified and thus that virus particles do not just sit inside latently infected cells. Rather, the viral DNA undergoes some biochemical change; whether or not this involves an integration event is under investigation.

We thank Drs H. Koprowski and N. Nathanson for their interest in this work, Gregory Dressler for technical assistance, Dr B. Roizman for *Bam*HI clones and L. Enquist for *Eco*RI clones. This work was supported by grant NS-11036 from the National Institute of Neurological and Communicative Disorders and Stroke, and by a grant NMS 851-C7 from the National Multiple Sclerosis Society. D.L.R. is the recipient of postdoctoral trainee grant NS 07180.

Received 20 September 1982; accepted 14 January 1983.

1. Stevens, J. G. & Cook, M. L. *Science* **173**, 843–845 (1971).
2. Cook, M. L. & Stevens, J. G. *J. gen. Virol.* **31**, 75–80 (1976).
3. Knotts, F. B., Cook, M. L. & Stevens, J. G. *J. exp. Med.* **138**, 740–744 (1973).
4. Cabera, C. V. *et al. Nature* **288**, 288–290 (1980).
5. Kastrukoff, L., Long, C., Doherty, P. C., Wroblewska, Z. & Koprowski, H. *Nature* **291**, 432–433 (1981).
6. Fraser, N. W., Lawrence, W. C., Wroblewska, Z., Gilden, D. H. & Koprowski, H. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6461–6465 (1981).
7. Locker, H. & Frenkel, N. *J. Virol.* **32**, 429–441 (1979).
8. Frenkel, N., Locker, H., Batterson, W., Hayward, G. S. & Roizman, B. *J. Virol.* **20**, 527–531 (1976).
9. Graham, B. J., Bengali, Z. & Vande Woude, G. F. *J. Virol.* **25**, 878–887 (1978).
10. Frenkel, N. in *Human Herpesviruses* (eds Nahmias, A. J., Dowdle, N. R. & Schinazi, R. F.) 91–120 (Elsevier, Amsterdam, 1981).
11. Peden, K., Mounts, P. & Hayward, G. S. *Cell* **31**, 71–80 (1982).
12. Maitland, N. J. *et al. J. gen. Virol.* **55**, 123–137 (1981).
13. Post, L. E., Conley, A. J., Mocarski, E. S. & Roizman, B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4201–4205 (1980).
14. Jacob, R. J., Morse, L. S. & Roizman, B. *J. Virol.* **29**, 448–457 (1979).
15. Lindahl, T. *et al. J. molec. Biol.* **102**, 511–530 (1976).
16. Herskowitz, I. & Hagen, D. A. *Rev. Genet.* **14**, 399–445 (1980).
17. Roizman, B. *Cell* **16**, 481–494 (1979).

Activation of acetylcholine receptors causes the partition of hydrophobic cations into postsynaptic membrane vesicles

C. Geoffrey Davis*, Shaul Hestrin†, Herbert Landahl‡, Adrienne S. Gordon*, Ivan Diamond* & Juan I. Korenbrot‡†

Departments of *Neurology, *Pharmacology, †Physiology and ‡Biochemistry, University of California, School of Medicine, San Francisco, California 94143, USA

In the continued presence of cholinergic ligands, the acetylcholine receptor–channel complex (AChR) in postsynaptic membranes undergoes a sequence of conformational changes. On addition of the ligand, the receptor rapidly changes from a closed channel to an open channel conformation¹, then slowly changes to a nonconducting state termed desensitization^{2,3}. The lifetime of the open channel conformation⁴ and the rate of desensitization⁵ are both dependent on the magnitude of the membrane potential, suggesting that the ligand-induced conformational changes in AChR may involve the movement of electrical charges within the membrane^{6,7}. Measurements of charge redistribution in AChR-containing membranes following ligand binding have not been reported. Recently, measurements of changes in the membrane partition coefficient of hydrophobic ions have been used to detect electrostatic changes in both biological⁸ and model^{9,10} membranes. We report here that cholinergic ligands induce changes in the partition coefficient of the hydrophobic cation tetraphenylphosphonium (TPP) into AChR-enriched membranes. The extent and time course of these changes in TPP partition coefficient are accounted for in a kinetic model. We conclude that TPP movement is a monitor of a molecular event which may be associated with the slow component of AChR desensitization.

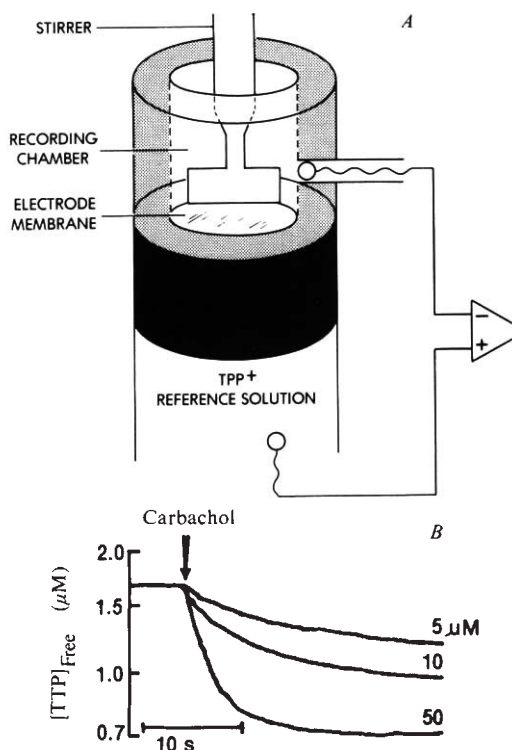


Fig. 1 A, Diagram of the TPP-selective membrane electrode, recording chamber and electrical configuration used to measure the concentration of TPP ions in the medium. B, Time course of change in TPP_{free} induced by addition of carbachol.

Methods: The membrane electrode was prepared by incorporating a mixture of TPP and tetraphenylboron (TPB) ions into a plasticized polyvinylchloride (PVC) matrix. Aqueous solutions of 50 ml of 10 mM TPP solution and 50 ml of 10 mM TPB solution were mixed and extracted overnight into 100 ml dichloroethane. Dichloroethane was then allowed to separate from the aqueous phase and rotoevaporated²⁶. The remaining powder was dissolved in 50 ml of peroxide-free tetrahydrofuran²⁷ and 3 ml of this solution was mixed with 0.4 g PVC, 1.5 ml dioctyl adipate, and 10 ml tetrahydrofuran. The resulting solution was poured over a 60 cm² glass surface. After solvent evaporation, a thin (approximately 100 μm) TPP-selective membrane electrode film was obtained. Small (8 mm diameter) circles of membrane electrode were cut and glued onto a cylindrical support machined from PVC (dark cylinder in diagram). An acrylic cylinder (shaded cylinder in diagram) was glued on top of the PVC cylinder. A cylindrical chamber of 300 μl volume was thus created, the floor of which was the membrane electrode. A motor-driven stirrer was used to minimize diffusion delays in the system. The recording chamber was mounted on the end of a U-shaped acrylic tube containing a reference solution with fixed TPP concentration. Electrical recordings of the potential across the TPP membrane electrode utilized a high input impedance (1,000 MΩ), low noise differential amplifier with a recording bandwidth of d.c. to 100 Hz (PARC, Model 113). The half-time of response of the electrode to a step change in TPP concentration was 150–300 ms. The selectivity of the electrode, measured by the mixed solution technique²⁸, had the following values: $K_{TPP/K^+} > 7.3 \times 10^{-7}$, $K_{TPP/Na^+} > 4.6 \times 10^{-7}$, $K_{TPP/Mg^{2+}} > 1.4 \times 10^{-6}$, $K_{TPP/Ca^{2+}} > 9.8 \times 10^{-7}$, $K_{TPP/Choline} > 7.1 \times 10^{-6}$, $K_{TPP/H^+} > 3.9 \times 10^{-2}$. In a typical experiment (B), 75 μl of AChR-enriched membranes prepared in 400 mM NaCl, 10 mM NaPO₄, pH 7.4, 1 mM EDTA was mixed in the recording chamber with an equal volume of 4 μM TPP in the same solution. The final concentrations were: AChR = 3.5 μM, TPP = 2 μM. In experiments in which ACh was used, 10 μM physostigmine was included in the incubation. The membranes were constantly stirred at room temperature until a stable baseline was achieved (about 1 min). The recorded baseline usually corresponded to approximately 1.7 μM free TPP, as indicated on the left. Ligand in a solution containing 1.7 μM TPP was then added in a volume of 15 μl. Both the rate and extent of the ligand-stimulated decline in the concentration of free TPP were dose-dependent. Addition of ligands in the absence of membranes had no effect.

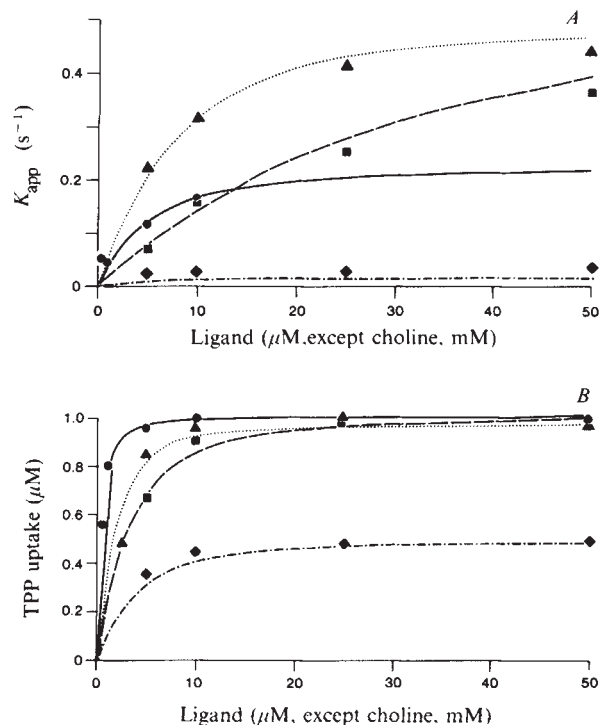


Fig. 2 A, Kinetics of TPP uptake. The points represent the value of k_{app} derived from the fit by a single exponential function to the rate of TPP uptake at the indicated ligand concentration: ■, carbachol; ▲, ACh; ●, choline; ◆, tubocurarine. Each data point is the average of 3–12 measurements. The range of values measured is within 20% of the average. The lines represent solutions of the model given in the text. The values of the kinetic constants K_0 , k_1 and k_2 used in drawing the lines are given in Table 1. B, Maximal TPP uptake. The points represent the average value of maximal TPP uptake at each ligand concentration. Symbols for the different ligands are as in A. In general, the range of the values measured is within 5% of the average. The lines represent solutions of the model given in the text.

AChR-enriched membrane vesicles were prepared from the electric organ of *Torpedo californica*¹¹. The AChR constitutes approximately 50% of the total protein in this preparation. The AChR specific activity in the membranes was about 3.5 nmol α-bungarotoxin (αBTx) binding sites per mg protein. The purified AChR-enriched membrane vesicles were washed in an electrolyte solution (see Fig. 1) and resuspended in the same solution at a concentration of 6.5–7.0 μM αBTx binding sites. Aliquots of 150 μl of the vesicle suspension containing 2 μM TPP were placed in a recording chamber which permitted rapid stirring and continuous measurement of the concentration of free TPP in the suspending medium. Free TPP was measured with a TPP-selective electrode (Fig. 1A). The AChR-enriched membrane vesicles steadily took up the free TPP from the medium over the first several minutes until the free TPP concentration attained a constant value. The sudden addition of cholinergic agonists resulted in a further rapid decrease of the free TPP concentration (Fig. 1B).

The time course of the ligand-induced TPP concentration decrease could be fitted by a single exponential function. The value of the rate constant (k_{app}) of this function is dependent upon both the choice of ligand and its concentration (Fig. 2A). The k_{app} increases as the ligand concentration increases up to a limiting value. The limiting value for acetylcholine (ACh) and carbachol is 0.42 ± 0.06 s⁻¹ (mean ± s.d.). For choline, k_{app} is 0.160 ± 0.052 s⁻¹ at 10 mM, the highest concentration tested. The k_{app} for the antagonist tubocurarine is much smaller and reaches a limiting value of 0.032 ± 0.010 s⁻¹. The k_{app} for all conditions tested reflects the rate of a molecular change in the

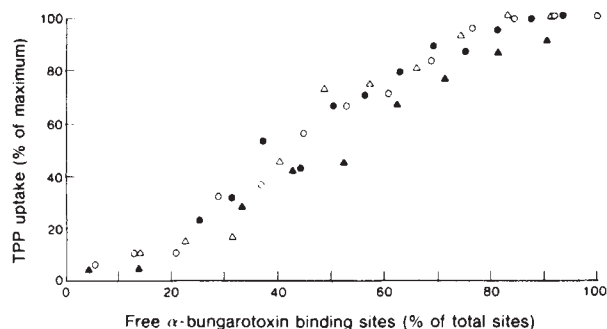


Fig. 3 Effect of partial receptor blockade on carbachol-induced uptake of TPP. In four different experiments (●, ○, ▲, △), membranes were preincubated with various amounts of α BTx for 60 min at 4 °C. Maximal TPP uptake (ordinate) in response to 100 μ M carbachol was then measured. In the graph, the amount of TPP uptake is shown as a function of free binding sites remaining after α BTx treatment (abscissa). The number of free sites was measured by an 125 I- α BTx binding assay.

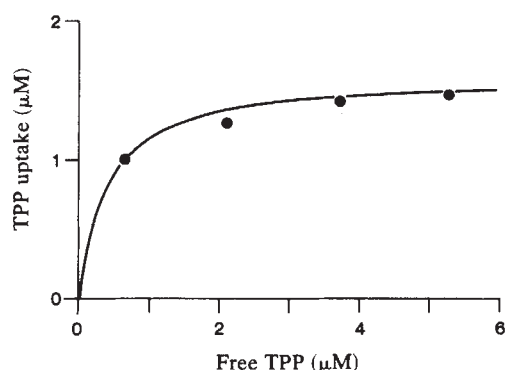


Fig. 4 Dependence of TPP uptake on TPP concentration. Membranes were equilibrated on the electrode with 2, 4, 6, or 8 μ M TPP. 100 μ M carbachol was then added and the maximal uptake of TPP was measured. Maximal TPP uptake (ordinate) is plotted versus the concentration of free TPP (abscissa). The results were analysed as a first-order reaction (line) and the plateau value for TPP uptake was 1.6 μ M.

membrane, as the addition of TPP to membrane vesicles first incubated with ligand results in a fast TPP uptake, the measured rate of which is limited by the time constant of the electrode.

The extent of the decrease in free TPP concentration is also proportional to agonist concentration (Fig. 2B). However, the limiting value of the concentration decrease was the same, 1 ± 0.05 μ M TPP, for all agonists tested and 0.49 ± 0.06 μ M for the antagonist tubocurare. The concentrations of ACh, carbachol, choline and tubocurare producing half-maximal TPP decrease were determined to be 2.5, 3.5, 300 and 4.5 μ M respectively.

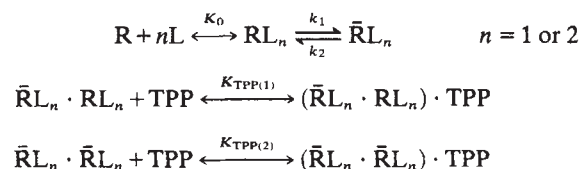
The distribution of TPP across membrane vesicles has previously been used as a probe of the membrane potential due to the unequal free concentrations of permeant ions between the inside and the outside of vesicles¹²⁻¹⁴. However, the distribution of TPP, like that of other hydrophobic ions, might also reflect hydrophobic or electrical interaction of the ion with the membrane itself^{8,9}. In the experimental system we describe here the ionic transmembrane potential should be 0 mV since the ionic solutions across the membrane are symmetric. Moreover, addition of 0.1% digitonin, a detergent concentration which permeabilizes AChR-enriched membranes to small ions, has no quantitative effect on the ligand-induced TPP uptake. Therefore, the TPP concentration decrease we report reflects the uptake of TPP by the membrane vesicles as a result of interaction of the hydrophobic ion with the membrane itself and not

as a result of changes in the macroscopic transmembrane potential.

Since the effects of cholinergic ligands could arise from their interaction with either the AChR or membrane-bound acetylcholinesterase, we tested the effect of carbachol on the partition of TPP into red cell ghosts which contain acetylcholinesterase but not AChR¹⁵. Addition of 25–250 μ M carbachol to red cell ghosts suspended at protein concentrations comparable to our standard assay had no effect on the free TPP concentration. Moreover, addition of carbachol to crude *Torpedo* membranes¹⁶ containing only 10–20% as much AChR as the membranes used in our standard assay resulted in the uptake of only about 10% as much TPP. These results indicate that the ligand-induced uptake of TPP into postsynaptic membranes is a result of ligand interaction with AChR.

If the TPP uptake is in fact a result of ligand binding to the AChR, it follows that it should be blocked by α BTx, a highly specific inhibitor of AChR. Additions of α BTx alone did not result in any TPP uptake. Figure 3 illustrates the extent of TPP uptake induced by a constant amount of carbachol added to membrane vesicles preincubated with various amounts of α BTx: the extent of ligand-induced TPP uptake is dependent on the fraction of sites available for ligand binding. The ligand-induced TPP uptake is fully reversible; addition of a 25 molar excess of α BTx to membranes first exposed to 100 μ M carbachol in the presence of TPP results in the release of the TPP specifically taken up. Thus, the ligand-induced TPP uptake is a fully reversible process which arises specifically from the excitation of the AChR molecule.

To analyse our data further, we developed the following model. This model fits both the kinetic and steady-state data we have obtained.



The model consists of three sequential reactions. In the first reaction, the receptor, R, binds ligand, L, to either of two equivalent sites with an equilibrium constant K_0 . The rate of formation of this complex is diffusion limited¹⁷. The ligand-receptor complex then spontaneously converts to a new state, $\bar{R}L_n$, with forward and backward rates k_1 and k_2 . The interaction between the $\bar{R}L_n$ state and TPP is modelled to be a first order reaction with equilibrium constant K_{TPP} . In the model, we consider the receptor molecules to exist as dimers in the membrane^{18,19}. The conversion of either or both of the monomers in the pair into the $\bar{R}$ state is sufficient to develop the increased affinity for TPP. This assumption is justified by the experimental fact that the stoichiometry of TPP uptake/AChR activated had a limiting value of 0.5 (see Fig. 4). We allowed the possibility that the TPP uptake resulting from only one monomer existing as $\bar{R}$ and the uptake resulting from both monomers existing in the $\bar{R}$ state may have different equilibrium constants, $K_{\text{TPP}(1)}$ and $K_{\text{TPP}(2)}$.

The solution of the kinetic equations generated by the model above were fitted to our experimental data and results are shown in Fig. 2 and Table 1. The values of the binding constant (K_0) for all four ligands determined by fitting the TPP data are within the range of values reported for this constant by direct measurements of others^{20,21}. The value of k_1 is similar to that reported for other slow processes associated with the AChR including changes in the binding constant for toxin^{22,23}, changes in intrinsic membrane fluorescence²⁴, changes in fluorescence of agonist probes²⁵, and the slow component of desensitization³.

In collaboration with D. Cafiso and W. Hubbell, we investigated the physical state of a nitroxide-labelled TPP analogue (their spin label I)⁹ in the AChR-enriched membranes following ligand addition. We found that this analogue is also taken up

Table 1 Parameters derived from TPP uptake using the sequential scheme described in the text

	K_0 (μM)	k_1 (s^{-1})	k_2 (s^{-1})
Carbachol	$55.7 \pm 4.5^*$	0.201 ± 0.006	0.042 ± 0.003
ACh	10.4 ± 1.7	0.176 ± 0.009	0.077 ± 0.008
Choline	$10,243 \pm 1,067$	0.081 ± 0.004	0.020 ± 0.001
Curare	3.2 ± 0.2	0.006 ± 0.0001	0.020 ± 0.0004

The solution of the kinetic equations generated by the model were fit to our experimental data by a least square regression routine to obtain estimates of the values of the kinetic parameters K_0 , k_1 and k_2 for each of the ligands tested. The fit was made to six measurements of TPP uptake from 0.5 to 10 s as well as to the steady state value for each ligand concentration. Therefore, the total number of data points fitted was 35 for carbachol and 28 for each of the other three ligands. The value of $K_{\text{TPP}(2)}$ was experimentally determined to be $0.4 \mu\text{M}$ from the data shown in Fig. 4. The best fit to the data was obtained under the assumption $K_{\text{TPP}(1)} = K_{\text{TPP}(2)}$, indicating that dimers with either one or two receptors in the $\bar{R}$ state are equally effective in taking up TPP. (Setting $K_{\text{TPP}(1)} = 0.8 \mu\text{M}$ increased the variance of the fit by 34%.)

* Root mean deviation of the difference between calculated and observed values.

by the membranes and that the line shape of its electron spin resonance spectrum did not change significantly upon partitioning. This suggests that the ligand-induced interaction of the hydrophobic ions with the membrane constituents is weak, as is also suggested by the high value of K_{TPP} .

The partition coefficient of TPP ions into a membrane, as any other ion, is a function of the difference in its free energies in the aqueous and the membrane phases. The free energy of hydrophobic ions consists of terms that arise from both their electrostatic and their hydrophobic interactions with the solvating medium. The physical nature of the TPP uptake reported here remains to be precisely determined. However, since the slow component of AChR desensitization is a voltage dependent process that occurs on a time scale commensurate with the measured rates of the TPP uptake we observe, it may be that the hydrophobic cation partition we have reported here reflects a change of electronegativity in the membrane associated with AChR desensitization.

We thank Ms T. Thomas for technical assistance. This work was supported by NIH grants NS 14742, EY 01586, and GM 27957, NSF grant BNS80-13074, and a grant and a fellowship (to S.H.) from the Muscular Dystrophy Association.

Received 10 August 1982; accepted 26 January 1983.

- Katz, B. & Miledi, R. *Nature* **226**, 962-963 (1970).
- Katz, B. & Thesleff, S. *J. Physiol., Lond.* **138**, 63-80 (1957).
- Sakmann, B., Patlak, J. & Neher, E. *Nature* **28**, 71-73 (1980).
- Magley, K. L. & Stevens, C. F. *J. Physiol., Lond.* **223**, 151-171 (1972).
- Scubon-Muliere, B. & Parsons, R. L. *J. gen. Physiol.* **71**, 285-299 (1978).
- Hodgkin, A. L. & Huxley, A. F. *J. Physiol., Lond.* **117**, 500-544 (1952).
- Stevens, C. F. *Biophys. J.* **22**, 295-306 (1978).
- Cafiso, D. S. & Hubbell, W. L. *Biophys. J.* **30**, 243-264 (1980).
- Cafiso, D. S. & Hubbell, W. L. *Biochemistry* **17**, 187-195 (1978).
- Andersen, O. J., Feldberg, S., Nakadomari, H., Levy, S. & McLaughlin, S. *Biophys. J.* **21**, 35-70 (1978).
- Gordon, A. S., Davis, C. G. & Diamond, I. *Proc. natn. Acad. Sci. U.S.A.* **74**, 263-267 (1977).
- Grinius, L. L. *et al. Biochim. biophys. Acta* **216**, 1-12 (1970).
- Ketterer, B., Neumcke, B. & Lauger, P. *J. Membrane Biol.* **5**, 225-245 (1971).
- Schuldiner, D. & Kaback, H. R. *Biochemistry* **14**, 5451-5461 (1975).
- Alles, G. A. & Howes, R. C. *J. biol. Chem.* **133**, 375-390 (1940).
- Hartig, P. R. & Raftery, M. A. *Biochemistry* **18**, 1146-1150 (1979).
- Boyd, N. D. & Cohen, J. B. *Biochemistry* **19**, 5353-5358 (1980).
- Chang, H. W. & Bock, E. *Biochemistry* **16**, 4513-4520 (1977).
- Hamilton, S. L., McLaughlin, M. & Karlin, A. *Biochemistry* **18**, 155-163 (1979).
- Boyd, N. D. & Cohen, J. B. *Biochemistry* **19**, 5344-5353 (1980).
- Kasai, M. & Changeux, J.-P. *J. Membrane Biol.* **6**, 1-23 (1971).
- Quast, U., Schimerlik, M., Lee, T., Witzemann, V., Blanchard, S. & Raftery, M. A. *Biochemistry* **17**, 2405-2415 (1978).
- Weber, M., David-Pfeuty, T. & Changeux, J.-P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3443-3447 (1975).
- Bonner, R., Barrantes, F. J. & Jovin, T. M. *Nature* **263**, 429-431 (1976).
- Heidmann, T. & Changeux, J.-P. *Eur. J. Biochem.* **94**, 255-279.
- Shimbo, T., Kamo, N., Karihara, K., Kobatake, Y. *Archs Biochem. Biophys.* **187**, 414-422 (1978).
- Gold, G. H. & Korenbrot, J. I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5557-5561 (1980).
- Moody, G. J., Oke, R. B. & Thomas, J. D. R. *Analyst* **95**, 910-918 (1970).

Structural homology of *Torpedo californica* acetylcholine receptor subunits

Masaharu Noda*, Hideo Takahashi*, Tsutomu Tanabe*, Mitsuyoshi Toyosato*, Sho Kikuyotani*, Yasuji Furutani*, Tadaaki Hirose†, Hideaki Takashima†, Seiichi Inayama†, Takashi Miyata‡ & Shosaku Numa*

* Department of Medical Chemistry, Kyoto University Faculty of Medicine, Kyoto 606, Japan

† Pharmaceutical Institute, Keio University School of Medicine, Tokyo 160, Japan

‡ Department of Biology, Kyushu University Faculty of Science, Fukuoka 812, Japan

The nicotinic acetylcholine receptor (AChR) from the electroplax of the ray *Torpedo californica* is composed of five subunits present in a molar stoichiometry of $\alpha_2\beta\gamma\delta$ (refs 1-3) and contains both the binding site for the neurotransmitter and the cation gating unit (reviewed in refs 4-6). We have recently elucidated the complete primary structures of the α -, β - and δ -subunit precursors of the *T. californica* AChR by cloning and sequencing cDNAs for these polypeptides^{7,8}. Here, we report the whole primary structure of the γ -subunit precursor of the AChR deduced from the nucleotide sequence of the cloned cDNA. Comparison of the amino acid sequences of the four subunits reveals marked homology among them. The close resemblance among the hydrophilicity profiles and predicted secondary structures of all the subunits suggests that these polypeptides are oriented in a pseudosymmetric fashion across the membrane. Each subunit contains four putative transmembrane segments that may be involved in the ionic channel. The transmembrane topology of the subunit molecules has also been inferred.

A library of cDNA clones derived from poly(A) RNA from an electric organ of *T. californica* was screened by hybridization with two kinds of synthetic oligodeoxyribonucleotide mixtures which represent all possible cDNA sequences corresponding to known amino acid sequences in the amino-terminal region of the γ -subunit³ (see Fig. 1 legend for experimental details). Two of the cDNA clones thus isolated (pACR γ 14 and pACR γ 16) were subjected to nucleotide sequence analysis⁹ according to the strategy indicated in Fig. 1a. Figure 1b shows the primary structure of the *T. californica* mRNA encoding the AChR γ -subunit precursor deduced from the cDNA sequence. By using the reading frame corresponding to the known amino-terminal 54-amino acid sequence of the mature γ -subunit³, the primary structure of the cryptic portion of its precursor was deduced (Fig. 1b). This precursor is composed of 506 amino acids, including a hydrophobic prepeptide of 17 amino acids which precedes the amino terminus of the mature γ -subunit. The molecular weight of the mature γ -subunit is calculated to be 56,279.

The alignment of the amino acid sequences of the α -, β -, γ - and δ -subunit precursors of the *T. californica* AChR (Fig. 2) shows conspicuous homology among them. Of the positions aligned, 19% and 20% are occupied by four and three identical residues, respectively, and 54% of the positions with three identical residues have a chemically similar residue to the fourth one. The hydrophilicity profiles¹⁰ as well as the secondary structures¹¹ predicted from the amino acid sequences of the four AChR subunits are similar, suggesting that all the subunits are oriented in a pseudosymmetric fashion with respect to the membrane. As described for the α -subunit⁷, the segments that can form secondary structures extend over a wide range of the subunit molecules, and each subunit has four strongly hydrophobic regions with secondary structure (corresponding to S5,

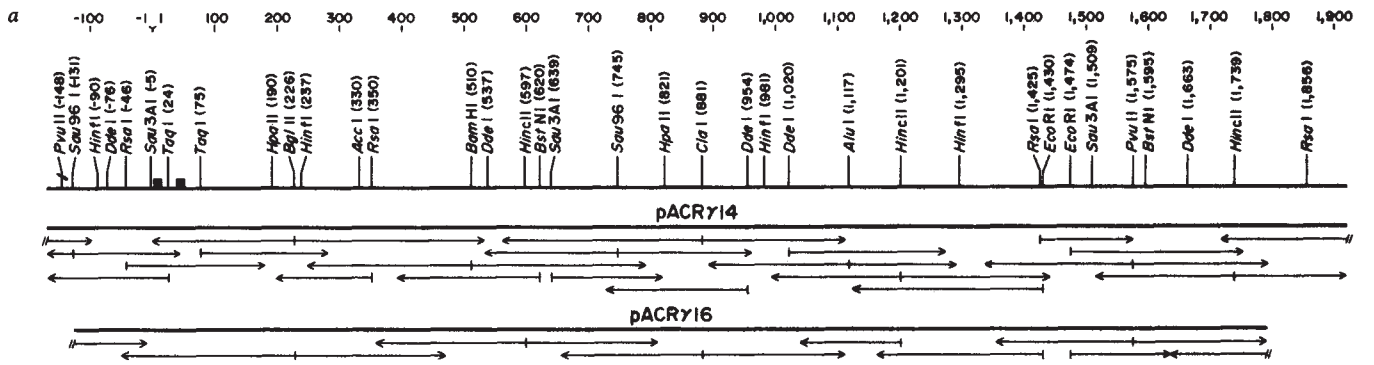


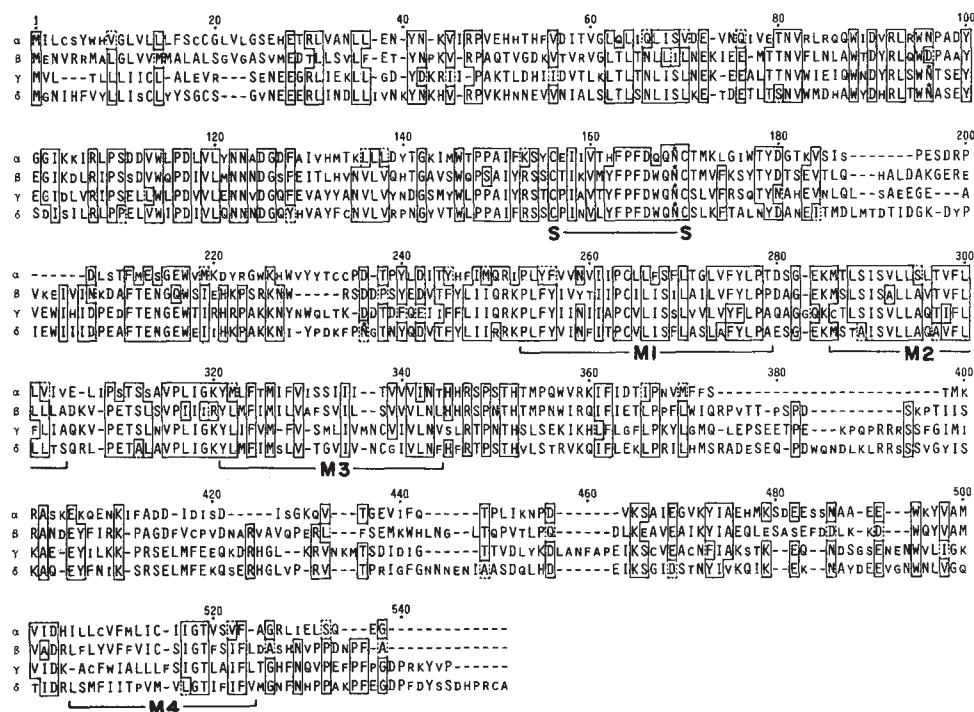
Fig. 1 Strategy for sequencing cloned cDNA encoding the γ -subunit precursor of the *T. californica* AChR (a) and the primary structure of the mRNA deduced from the cDNA sequence (b). The cDNA library used (constructed with the plasmid DNA vector of Okayama and Berg³⁵) was the same as that screened for cDNA encoding the α -, β - and δ -subunit precursors^{7,8}. Two kinds of oligodeoxyribonucleotide mixtures, synthesized by the triester method on a solid support³⁶, were used as hybridization probes. One of them was a mixture of 16 tetradecamers, 5'-

CC^TTC^TTC^TAT^TTC^T-3', which corresponds to the sequence Glu-Asn-Glu-Glu-Gly (residues 1-5 in b), excluding the third nucleotide of the Gly codon. The other was a mixture of 32 tetradecamers, 5'-TT^ATC^AGT^ATCNCC-3' (synthesized as two

pools containing A/T or G/C at the 12th position where N stands for a mixture of A, T, G and C), which corresponds to the sequence Gly-Asp-Tyr-Asp-Lys (residues 13-17 in b), excluding the third nucleotide of the Lys codon. Hybridization was carried out at 37°C with the probes labelled with ³²P at the 5' end. Thirty-four clones that hybridized with both probes were isolated from ~3 × 10⁵ transformants and classified by restriction endonuclease mapping into two groups of about an equal population, which had a different 3' terminus (see below). One clone from each group (pACR γ 14 and pACR γ 16), which harboured the apparently largest cDNA insert in the respective groups, was subjected to DNA sequencing by the procedure of Maxam and Gilbert⁹. The restriction map in a displays only relevant restriction endonuclease sites, which are identified by numbers indicating the 5'-terminal nucleotide generated by cleavage (for the nucleotide numbers, see b). The poly(dA)·poly(dT) tract and poly(dG)·poly(dC) tail are not included in the restriction map. The sequences corresponding to the hybridization probes used are indicated by closed boxes. The direction and extent of sequence determinations are shown by horizontal arrows under each clone used; the sites of 5'-end labelling are indicated by short vertical lines at the end of arrows. The slash marks at the end of arrows mean that the site of 5'-end labelling was the BstNI site (left marks) or the PvuII site (right marks) located on the vector DNA^{35,37}. Further details of the sequencing procedures have been described previously³⁸. In b, the nucleotide sequence of the mRNA was deduced from those of the cDNA inserts in clones pACR γ 14 and pACR γ 16. No

nucleotide difference was observed between the sequences determined for the two clones. Nucleotide residues are numbered in the 5' to 3' direction, beginning with the first residue of the codon specifying the amino-terminal glutamic acid of the mature γ -subunit, and the nucleotides on the 5' side of residue 1 are indicated by negative numbers. The translational initiation site was assigned to the AUG codon represented by nucleotide residues -51 to -49 because this is the first AUG triplet that appears downstream from the nonsense codons (residues -120 to -118, -123 to -121 and -138 to -136) found in-frame. The predicted amino acid sequence is displayed above the nucleotide sequence. Amino acid residues are numbered beginning with the amino-terminal glutamic acid of the mature γ -subunit, and the preceding residues are indicated by negative numbers. Clone pACR γ 16 carries an mRNA sequence that is polyadenylated after residue 1,790 (indicated by an arrow), 14 nucleotides downstream from the AAUAAA sequence³⁹ (residues 1,772-1,777). No known polyadenylation signal^{39,40} is present upstream of the polyadenylation site (after residue 1,921; indicated by an arrow) for the mRNA sequence harboured by clone pACR γ 14, although some A+U-rich sequences are found⁴¹; four other clones of the same group also lack a known polyadenylation signal. There are two additional AAUAAA sequences (residues 1,551-1,556 and 1,555-1,560) which partially overlap. Blot hybridization analysis^{42,43} of electroplax poly(A) RNA with the PvuII(-148)-PvuII(+1,575) fragment (labelled by nick-translation⁴⁴ with [α -³²P]dCTP) exhibited an apparently single hybridization-positive band corresponding to a size of ~2,300 nucleotides. The sequence of nucleotide residues -82 to 162 is in complete agreement with the partial cDNA sequence reported by Ballivet *et al.*⁴⁵.

Fig. 2 Comparison of the amino acid sequences of the AChR subunit precursors. The amino acid sequences of the α - (top), β - (second row), γ - (third row) and δ -subunit precursors (bottom) are aligned; the sequence data for the α -, β - and δ -subunit precursors have been taken from refs 7 and 8. The one-letter amino acid notation is used. Sets of identical residues and residues considered to be favoured amino acid substitutions are represented by large letters, and the remaining residues by small letters. Sets of three or four identical residues at one aligned position are enclosed with solid lines, and the fourth residues regarded as favoured substitutions at the same position are enclosed with dotted lines. Favoured amino acid substitutions are defined as pairs of residues belonging to one of the following groups: S, T, P, A and G; N, D, E and Q; H, R and K; M, I, L and V; F, Y and W (ref. 46). Gaps (-) have been inserted to achieve maximum homology. The positions in the aligned sequences including gaps are numbered beginning with that of the initiative methionine. The putative disulphide bridge proposed as being in close proximity to the acetylcholine binding site on the α -subunit^{7,8} is shown (S—S); the other subunits may also contain a corresponding disulphide bridge. The possibility that one of the remaining two cysteine residues (positions 233 and 234) in the extracellular region of the α -subunit participates in the disulphide bridge near the acetylcholine binding site⁷ cannot be excluded. The putative transmembrane segments (M1–M4 in Fig. 3) are also shown. The asparagine residues as possible sites of glycosylation are indicated by asterisks, except for the two sites in the γ -subunit assigned to the cytoplasmic side of the membrane.



S6, S7 and S10 for the α -subunit⁷), which may represent transmembrane segments or may be involved in inter-subunit interaction⁷. These hydrophobic regions of the four subunits contain a continuous stretch of 19–27 uncharged amino acid residues including many nonpolar residues (68–93%) (Fig. 3). These segments, designated M1, M2, M3 and M4, are bounded by charged residues. Because such a sequence is characteristic of transmembrane protein segments^{12–17}, the segments M1–M4 most probably traverse the membrane. In view of the lengths of these segments and by analogy with the transmembrane segments of bacteriorhodopsin¹⁸ and glycophorin A (refs 19, 20), it seems reasonable to assume that the segments M1–M4 form α -helical structures. The segment M1 contains three conserved proline residues (positions 253, 263 and 278) which introduce bends with an angle of 20° in the helix axis. (Note that the numbers indicating positions in the aligned amino acid sequences differ from those of amino acid residues given in Fig. 1b and refs 7 and 8.)

It is tempting to assume that these transmembrane segments are involved in the cation channel of the AChR. Interestingly, the polar side chains present in these segments are clustered largely on one side of an α -helix (Fig. 3). The endplate channel, which is permeable to monovalent and divalent cations and non-electrolytes but discriminates against anions, is considered to be a water-filled neutral pore without high-field strength sites inside²¹. The uncharged nature of the segments M1–M4 is consistent with this concept and may underlie the weak selectivity of the ionic channel of the AChR. In view of the diameter of the central pore being 6–25 Å (refs 21–25), one or two helical segments from each subunit together would suffice to construct the wall of the ionic channel, and the remaining helical segments might interact with those forming the channel. In this context, it seems more likely that the segments M1 and M2, which are highly conserved among the four subunits, are more directly involved in the ionic channel. The segment M1 contains three proline residues (see above) and one cysteine residue (position 264) which are conserved in all four subunits. This is not typical of transmembrane segments acting simply

as a lipophilic membrane anchor¹⁷, but favours the view that the segment M1 carries a specific function. The cysteine residue in the segment M1 is the only cysteine residue conserved in all subunits other than the two cysteine residues (positions 156 and 170) proposed^{7,8} as taking part in the disulphide bridge near the acetylcholine binding site on the α -subunit⁵.

Studies on selective proteolysis of the AChR subunits have indicated that the amino-terminal two-thirds of the subunit molecules are resistant to proteolysis and are glycosylated^{26–28}. In fact, this region contains the single possible *N*-glycosidic linkage site (position 169) in the α - and β -subunits^{7,8}, all of the three possible glycosylation sites (positions 96, 169 and 236) in the δ -subunit⁸ and two of the four possible glycosylation sites (positions 96, 169, 343 and 486) in the γ -subunit. Thus, the amino-terminal half of the mature subunit molecules preceding the transmembrane segment M1 is most probably exposed on the extracellular surface of the membrane, forming a rigid protein core due to secondary, tertiary and quaternary structures (see above and ref. 7). There is also evidence indicating that the carboxy-terminal region of the δ -subunit carries the disulphide bond responsible for the dimerization of the receptor²⁷ and that the segment linking the δ -subunits of adjacent receptor molecules is located outside the membrane²⁹. The cysteine residue (position 550) next to the carboxyl terminus of the δ -subunit is the only cysteine that meets these requirements. On the basis of these findings, the orientation of the transmembrane segments M1–M4 has been assigned as shown in Fig. 3. Thus, the region between segments M3 and M4, constituting about one-quarter of the mature subunit molecules, is located on the cytoplasmic side of the membrane. The proposed transmembrane topology of the AChR molecule is consistent with the estimations hitherto made by physical and morphological approaches^{23,30,31}.

We have previously predicted several possible candidates for antigenic determinants on the α -subunit⁷. The proposed transmembrane topology of the subunit molecules eliminates some of these sites because they are assigned to the cytoplasmic side of the membrane. The peptide segment Glu-Ser-Asp-Arg-Pro-

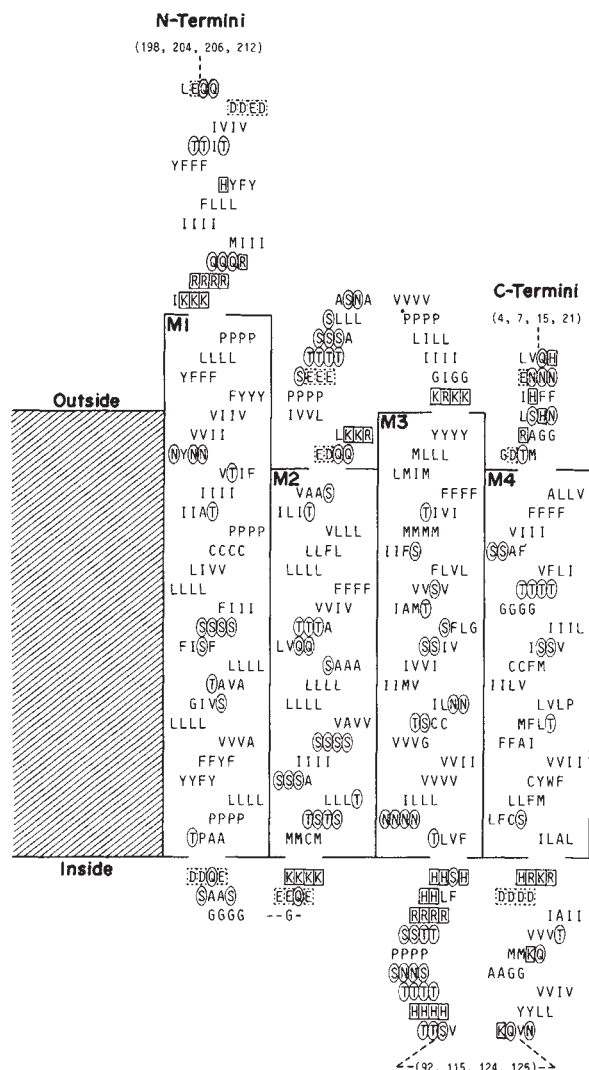


Fig. 3 Proposed orientation of the AChR subunit polypeptides across the membrane. The amino acid sequences of the four putative transmembrane segments (designated as M1-M4 in the direction from amino terminus to carboxyl terminus and boxed) and the adjacent regions are shown by the one-letter notation. Sets of four letters indicate the residues in the aligned α -, β -, γ - and δ -subunit sequences, respectively, from left to right; gaps (-) have been removed, except at position 283, to permit ready identification of the side chains extending towards one side of an α -helix. Sets of four numbers in parentheses indicate the sum of residues in the mature α -, β -, γ - and δ -subunit sequences, respectively (from left to right), that either extend beyond the ends of the displayed sequences or are inserted between them. Uncharged polar residues⁴⁷ (Ser, Thr, Asn and Gln) are enclosed with solid circles, positively charged residues with solid squares, and negatively charged residues with dotted squares. The termini of the proposed transmembrane segments have been tentatively assigned to the ends of a region composed exclusively of uncharged residues, so that the lengths of the boxed segments do not necessarily represent the real lengths of the segments embedded in the lipid hydrocarbon portion (shaded) of the membrane.

Asp (positions 196-207 with gaps) on the α -subunit seems to be the most likely candidate for the 'main immunogenic region' (refs 32-34), which is distinct from but may be near the acetylcholine binding site, although other regions on the extracellular surface of the α -subunit corresponding to local upspikes of the hydrophilicity profile⁷ cannot be excluded.

The amino acid sequence homology observed among the four subunit precursors of the AChR (Fig. 2) strongly suggests that the genes encoding them descended from a single common ancestor by gene duplications. A phylogenetic tree representing

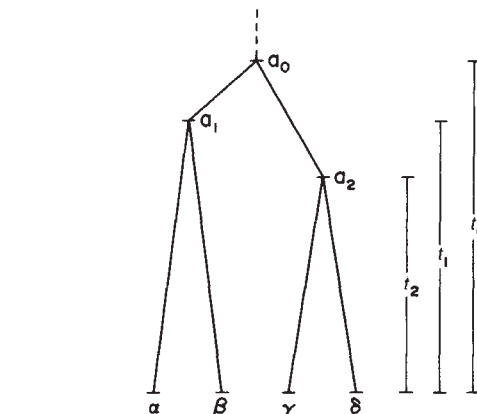


Fig. 4 Evolutionary relationships among the α -, β -, γ - and δ -subunit precursors of the *T. californica* AChR. By comparing the aligned amino acid sequences (Fig. 2), the amino acid difference per residue (K) was calculated for each pair of the mature subunits; K is defined as $K = M/N$, where M is the number of positions at which different amino acids are found between the sequences compared, and N the total number of positions; gaps were counted as one substitution regardless of their length. The K values obtained and their corrections for multiple substitutions⁴⁸ (in parentheses) are as follows: $K(\alpha/\beta) = 0.57$ (0.85); $K(\alpha/\gamma) = 0.64$ (1.03); $K(\alpha/\delta) = 0.63$ (0.99); $K(\beta/\gamma) = 0.58$ (0.87); $K(\beta/\delta) = 0.59$ (0.88); $K(\gamma/\delta) = 0.43$ (0.57). Note that the K values for the pairs α/γ and α/δ are larger than those for the pairs β/γ and β/δ . This implies that the rate of evolution differs between the lineages α_1 - α and α_1 - β . Assuming a constant rate of evolution (v) over all lineages, except for the lineage α_1 - α with a different rate of evolution (v_α), the phylogenetic tree was constructed as described previously⁴⁹. The relative dates of branching (t_0 , t_1 and t_2 in years) and the rates of evolution (per residue per year) are as follows: $t_1 = 0.82t_0$; $t_2 = 0.65t_0$; $v = 0.44t_0^{-1}$; $v_\alpha = 0.60t_0^{-1}$.

the evolutionary relationships among the four subunits is shown in Fig. 4.

We thank Drs Paul Berg and Hiroto Okayama for providing us with their high-efficiency cloning system. This investigation was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases.

Note added in proof: Claudio *et al.*⁵⁰ have also determined the nucleotide sequence of cloned cDNA for the γ -subunit precursor of the *T. californica* AChR.

Received 6 December 1982; accepted 3 February 1983.

- Reynolds, J. & Karlin, A. *Biochemistry* **17**, 2035-2038 (1978).
- Lindstrom, J., Merlie, J. & Yogeewaran, G. *Biochemistry* **18**, 4465-4470 (1979).
- Rafferty, M. A., Hunkapiller, M. W., Strader, C. D. & Hood, L. E. *Science* **208**, 1454-1457 (1980).
- Heidmann, T. & Changeux, J. P. *Rev. Biochem.* **47**, 317-357 (1978).
- Karlin, A. in *Cell Surface Reviews* Vol. 6 (eds Cotman, C. W., Poste, G. & Nicolson, G. L.), 191-260 (North-Holland, Amsterdam, 1980).
- Conti-Tronconi, B. M. & Rafferty, M. A. *Rev. Biochem.* **51**, 491-530 (1982).
- Noda, M. *et al. Nature* **299**, 793-797 (1982).
- Noda, M. *et al. Nature* **301**, 251-255 (1983).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824-3828 (1981).
- Chou, P. Y. & Fasman, G. D. *Rev. Biochem.* **47**, 251-276 (1978).
- Robb, R. J., Terhorst, C. & Strominger, J. L. *J. biol. Chem.* **253**, 5319-5324 (1978).
- Engelman, D. M., Henderson, R., McLachlan, A. D. & Wallace, B. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2023-2027 (1980).
- Rogers, J. *et al. Cell* **20**, 303-312 (1980).
- Uehara, H., Coligan, J. E. & Nathenson, S. G. *Biochemistry* **20**, 5940-5945 (1981).
- Ross, A. H., Radhakrishnan, R., Robson, R. J. & Khorana, H. G. *J. biol. Chem.* **257**, 4152-4161 (1982).
- Rice, C. M., Bell, J. R., Hunkapiller, M. W., Strauss, E. G. & Strauss, J. H. *J. molec. Biol.* **154**, 355-378 (1982).
- Henderson, R. & Unwin, P. N. T. *Nature* **257**, 28-32 (1975).
- Schulte, T. H. & Marchesi, V. T. *Biochemistry* **18**, 275-280 (1979).
- Egmond, M. R., Williams, R. J. P., Welsh, E. J. & Rees, D. A. *Eur. J. Biochem.* **97**, 73-83 (1979).
- Adams, D. J., Dwyer, T. M. & Hille, B. *J. gen. Physiol.* **75**, 493-510 (1980).
- Maeno, T., Edwards, C. & Anraku, M. *J. Neurobiol.* **8**, 173-184 (1977).
- Klymkowsky, M. W. & Stroud, R. M. *J. molec. Biol.* **128**, 319-334 (1979).
- Cartaud, J. & Benedetti, E. L. *FEBS Lett.* **33**, 109-113 (1973).
- Nickel, E. & Potter, L. T. *Brain Res.* **57**, 508-517 (1973).
- Wennogle, L. P. & Changeux, J. P. *Eur. J. Biochem.* **106**, 381-393 (1980).
- Wennogle, L. P., Oswald, R., Saitoh, T. & Changeux, J. P. *Biochemistry* **20**, 2492-2497 (1981).

28. Strader, C. D. & Raftery, M. A. *Proc. natn. Acad. Sci. U. S. A.* **77**, 5807-5811 (1980).
29. Conti-Tronconi, B. M., Dunn, S. M. J. & Raftery, M. A. *Biochemistry* **21**, 893-899 (1982).
30. Wise, D. S., Karlin, A. & Schoenborn, B. P. *Biophys. J.* **28**, 473-496 (1979).
31. Zingsheim, H. P., Neugebauer, D. C., Frank, J., Hänicke, W. & Barrantes, F. J. *EMBO J.* **1**, 541-547 (1982).
32. Tzartos, S. J., Rand, D. E., Einarson, B. L. & Lindstrom, J. M. *J. biol. Chem.* **256**, 8635-8645 (1981).
33. Tzartos, S. J., Seybold, M. E. & Lindstrom, J. M. *Proc. natn. Acad. Sci. U. S. A.* **79**, 188-192 (1982).
34. Gullick, W. J., Tzartos, S. & Lindstrom, J. *Biochemistry* **20**, 2173-2180 (1981).
35. Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161-170 (1982).
36. Ito, H., Ike, Y., Ikuta, S. & Itakura, K. *Nucleic Acids Res.* **10**, 1755-1769 (1982).
37. Reddy, V. B. *et al. Science* **200**, 494-502 (1978).
38. Nakanishi, S. *et al. Nature* **278**, 423-427 (1979).
39. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211-214 (1976).
40. Goeddel, D. V. *et al. Nature* **290**, 20-26 (1981).
41. Setzer, D. R., McGrogan, M. & Schimke, R. T. *J. biol. Chem.* **257**, 5143-5147 (1982).
42. McMaster, G. K. & Carmichael, G. G. *Proc. natn. Acad. Sci. U. S. A.* **74**, 4835-4838 (1977).
43. Alwine, J. C., Kemp, D. J., Stark, G. R. *Proc. natn. Acad. Sci. U. S. A.* **74**, 5350-5354 (1977).
44. Weinstein, R., Sweet, R., Weiss, M., Cedar, H. & Axel, R. *Proc. natn. Acad. Sci. U. S. A.* **75**, 1299-1303 (1978).
45. Ballivet, M., Patrick, J., Lee, J. & Heinemann, S. *Proc. natn. Acad. Sci. U. S. A.* **79**, 4466-4470 (1982).
46. Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. in *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3, 345-352 (National Biomedical Research Foundation, Silver Spring, Maryland, 1978).
47. Capaldi, R. A. & Vanderkooi, G. *Proc. natn. Acad. Sci. U. S. A.* **69**, 930-932 (1972).
48. Kimura, M. & Ohta, T. *J. molec. Evol.* **2**, 87-90 (1972).
49. Miyata, T. & Hayashida, H. *Nature* **295**, 165-168 (1982).
50. Claudio, T., Ballivet, M., Patrick, J. & Heinemann, S. *Proc. natn. Acad. Sci. U. S. A.* **80**, 1115 (1983).

Direct evidence for microfilament-mediated capping of surface receptors on crawling fibroblasts

Julian P. Heath

MRC Cell Biophysics Unit, Department of Biophysics,
King's College, 26-29 Drury Lane, London WC2B 5RL, UK

On moving fibroblasts, cell-surface receptors cross-linked by antibodies or lectins are cleared centripetally from regions of lamellar cytoplasm and collect as a cap over the perinuclear region¹⁻⁴. Current theories of the mechanism of receptor redistribution on cultured cells variously implicate membrane flow⁵⁻⁷, lipid flow⁸, surface waves^{9,10} and linkage to the cytoskeleton¹¹⁻¹³. The last, the anchorage model¹³, is based on observations that ligand-induced clusters of receptors on a variety of cell types either attach to actin¹⁴ or align over structures containing actin¹⁵⁻¹⁷, myosin^{15,16} and α -actinin¹⁸. I show here that the capping of antibody receptors on crawling chick embryo fibroblasts is highly coordinated with the apparent centripetal movements of arcs¹⁹⁻²², which are part of a dorsal cortical actin-microfilament sheath (DCMS). This phenomenon can be directly observed in living cells. The data support the anchorage model of membrane receptor mobility and suggest that there is a continuous flow of actin associated with fibroblast locomotion.

During fibroblast locomotion, protrusion of the leading lamella often leads to the formation of arcs¹⁹⁻²², which develop close to the lamellar margin and then, as judged by light microscopy, move centripetally through the lamella before disappearing in the perinuclear region. Chick embryo fibroblasts¹⁹⁻²¹ contain a 100-nm-thick cortical sheath of circumferentially arranged actin filaments which lie immediately beneath the dorsal surface of the lamella (Fig. 1a). Arcs are regions of this sheath where the filaments are more densely packed and so visible in the light microscope (Fig. 1c, e). Arcs and the DCMS stain with anti-actin and anti-myosin antibodies (Fig. 1b-e). About 34% of the fan-shaped and circular cells in respreading cultures of chick fibroblasts have one or more arcs moving centripetally through their lamellae²¹. Arc movement occurs simultaneously and coordinately with the transport of particles on the lamella^{5,21}, suggesting that the DCMS may control the mobility of surface receptors and their ligands.

The relationships between the movement of arcs and the distribution of cell-surface receptors were examined using an

antiserum from a rabbit immunized with cultured chick embryo sympathetic neurones (CSN). Anti-CSN specifically stains the surfaces of chick fibroblasts, epithelial cells and neural cells, behaving as a monovalent ligand because cross-linking by incubation in a second antibody is required before patching and capping of the anti-CSN receptors occurs¹¹.

Figure 1f shows the normal distribution of anti-CSN receptors on the surface of chick heart fibroblasts. Cells were incubated in anti-CSN at 37 °C, fixed and the antibody traced using rhodamine-conjugated goat anti-rabbit IgG (RhGARIG). All fibroblasts have a uniform, punctate distribution of anti-CSN receptors that is unperturbed by any underlying cytoplasmic structures, including arcs. Some cells were labelled with anti-CSN and their movements were followed for up to 60 min to see what effects locomotion and arc formation had on receptor distribution. Figure 1g-i illustrate a typical experiment. During a period of 22 min after anti-CSN labelling, the cell formed five arcs (see, for example, Fig. 1g, h) and moved forward a distance of 30 μ m. At the end of this period the cell was fixed and labelled with RhGARIG (Fig. 1i). In this and other cells examined in this manner there was no redistribution of anti-CSN receptors as a result of either arc formation or locomotion. This second finding confirms work by Middleton²³ on rat fibroblasts and argues against the membrane flow model of cell locomotion and capping^{6,7}, as anti-CSN is still present on the newly protruded regions of the lamella.

When living fibroblasts are incubated in anti-CSN followed by RhGARIG at 37 °C, the receptors aggregate into clusters, or patches, about 1-2 μ m wide (Fig. 1j). On non-motile, polygonal cells (see, for example, Fig. 1j) the patches are often found aligned over stress fibres, as has been previously reported¹⁵⁻¹⁷, but on moving or respreading fibroblasts the receptor patches are capped (Fig. 1k, o). Patches are cleared only from the leading lamella and from those other regions at the cell margin that are engaged in protrusion; there is no clearance of patches on the surfaces of stationary cells or on the retracted regions of moving cells (Fig. 1k, l), supporting the widely held view that capping and cell movement are related processes perhaps sharing common mechanisms^{2,11}.

Using an SIT video camera and video recorder, high resolution recordings were obtained of anti-CSN capping on moving cells. It was discovered that receptor patches are not cleared sequentially but as a unit, with all patches retaining their relative positions on the lamella as they are swept rearwards to the nucleus. The boundary between the retreating patches and the cleared lamella is always well defined and follows the contours of the lamellar margin (Fig. 1n). On reaching the perinuclear zone the patches slow down and stop moving and form a large fluorescent cap (Fig. 1o). Patches move centripetally at 1-3 μ m min⁻¹ which is comparable with the speed of other centripetally moving structures in or on the lamella such as particles⁵, ruffles²⁴, blebs⁶ and arcs^{20,21}. Furthermore, like particles²⁵, patches are also cleared from the ventral cell surface (Fig. 1k) but their movement generally ceases around 5 μ m from the cell margin, probably because they become trapped between cell and substratum; many, if not all, of these patches are left behind on the substratum as a cell moves forward.

When anti-CSN receptors are cross-linked on cells that contain centripetally moving arcs, a striking pattern of fluorescence is observed. By staining cells with RhGARIG for different periods it was noted that after 10 or 30 s of incubation in RhGARIG, newly formed patches were uniformly distributed over the lamella surface; but at 1 min after cross-linking, the density of patches was three to five times greater over arcs than elsewhere on the lamella (Fig. 1n,p). Although earlier studies¹⁵⁻¹⁷ showed that patches tend to congregate over microfilament bundles, the speed at which this association occurs seems not to have been previously recorded. During the capping process on these cells, there is precise registration at all times of the regions of increased patch density and the underlying arcs. If an arc formed before cross-linking with RhGARIG, patches distal to the arc followed the arc rearwards,

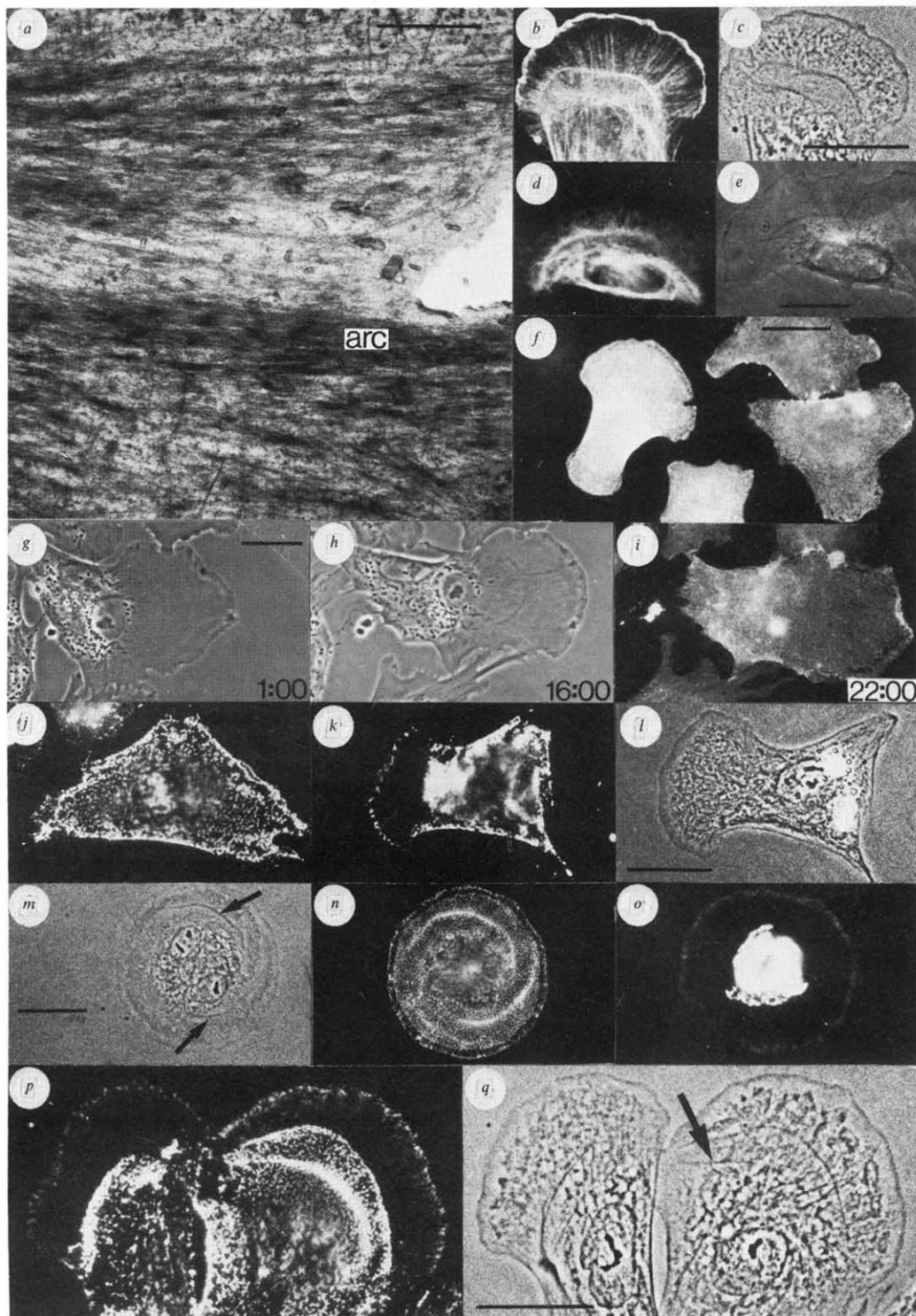


Fig. 1 The relationships between locomotion, the centripetal movement of arcs and anti-CSN distribution on chick embryo fibroblasts. Cells were grown from 8-day chick embryo heart explants cultured in plastic flasks under medium 199 plus fetal bovine serum (10%), glutamine (2 mM), penicillin (100 IU ml⁻¹) and streptomycin (100 µg ml⁻¹) (M199). After 48 h the cells were trypsinized, seeded onto glass coverslips and cultured for 1–3 h before use. Methanol-fixed or living cells were incubated in anti-actin, anti-myosin or anti-CSN antisera followed by RhGARig; all incubations were at 37 °C for 2–5 min using 1:50 dilutions in M199. The cells were examined on a Zeiss microscope and photographed on Tri-X or HP5 film; receptor redistributions on living cells were recorded on videotape using a RCA SIT video camera. For further details of techniques and electron microscopy, see ref. 21. Scale bars: *a*, 1 µm; *b–q*, 20 µm. *a*, Electron micrograph of a horizontal section through a fibroblast lamella showing the dorsal cortical sheath of circumferentially oriented actin microfilaments, interspersed with dense bodies, and an arc. On the right the section passes out of the dorsal cell surface. *b–e*, Paired immunofluorescence and phase contrast micrographs of arc-containing fibroblasts stained with anti-actin (*b*, *c*) and anti-myosin (*d*, *e*). *f*, The normal, uniform distribution of anti-CSN receptors on the surfaces of moving chick fibroblasts. *g–i*, Time sequence showing absence of anti-CSN receptor redistribution on a moving cell during arc formation. This cell was incubated with anti-CSN, returned to normal medium at 37 °C for 22 min (*g*, *h*) and then fixed and labelled with RhGARig (*i*). *j*, Patches of anti-CSN receptors on a stationary living fibroblast. *k*, *l*, Clearance of patched receptors from the lamellar regions of a moving fibroblast. The anterior ring of fluorescence is on the ventral cell surface. *m–o*, Capping of receptors on circular respreading cells. *m*, *n* Show the amount of receptor clearance at 2 min after cross-linking with RhGARig; note the increased density of patched receptors over the two arcs (arrows). *o* Shows a similar cell which has capped, 20 min after cross-linking. *p*, *q*, Coordinate centripetal movement of arcs and receptor patches on two crawling fibroblasts. In the right-hand cell, a second arc (closer to the nucleus and partly out of focus, arrowed) coincides with a dense band of patches.

maintaining their positions relative to the arc until reaching the perinuclear region (Fig. 1n). If an arc formed at the time of or close to patching, the centripetal movement of the arc seemed to sweep the lamella clear of patches (Fig. 1p, q), but in these cases also, patches closer to the nucleus retained their position relative to the arc during capping. The ability of these cells to form arcs, and the frequency of arc formation were no different in anti-CSN-labelled cells than in unlabelled controls; moreover, capped fibroblasts continued to form arcs as normal.

This study provides new evidence, directly observed in living cells, for the involvement of cytoplasmic actin microfilaments in the mobility of cell-surface receptors. Submembranous microfilaments have long been considered likely candidates for moving cross-linked receptors through the membrane and into a cap¹¹, but earlier studies that showed a coordinate redistribution of receptors and cytoskeletal proteins^{17,18,26-29} are open to the criticism that the receptors have moved to regions where these proteins are concentrated. In this study, however, I have shown that microfilament rearrangements can occur continuously in moving fibroblasts with no apparent effect on the distribution of surface receptors until they have been cross-linked. Only then do the receptors redistribute, coordinately with the microfilaments, into a cap. The crucial observations that the patched anti-CSN receptors move as a unit and not individually and that they are more dense over arcs, where the DCMS filaments are most concentrated, implies that the patches are associated with the submembranous DCMS.

Given this association, how are the receptors moved centripetally? It seems improbable that arcs are longitudinal or compression waves moving through a stationary DCMS, as postulated by Sorzano and Bell²², or that arcs create a transverse wave in the membrane and displace the receptors by a surfboarding mechanism^{9,10}, for such mechanisms would be expected to clear the patches sequentially. Moreover, I have shown that capping does not only occur in arc-forming cells.

I suggest that a simple hypothesis to explain these findings is that there is a continuous centripetal flow of DCMS material beneath the surfaces of the lamellae of moving and spreading cells. DCMS movement ceases when the cells are stationary, hence non-motile cells do not cap. The DCMS would be assembled at the lamellar margin and disassembled near the nucleus where the movement of patches, particles and arcs stops. Thus, there would need to be a compensatory forward flow of material through the lamella³⁰. The mechanisms of assembly of the DCMS at the lamellar margin are envisaged as intimately related to those that protrude the lamella forward, which explains the correlation between arc formation and increased protrusive activity²¹. The circumferential orientation of the DCMS filaments, and the presence of myosin, suggest that movement could be generated by a sliding filament mechanism which would tend to move the filaments centripetally. Any membrane receptors that as a result of cross-linking, formed an attachment to or were constrained to move with the DCMS would be capped. The observation that patches are also cleared from the ventral surfaces of moving fibroblasts suggests that the centripetal flow of microfilaments may not be confined to the DCMS³¹.

I thank Dennis Bray and Clive Thomas for the gift of anti-CSN antiserum, and Graham Dunn and Adam Middleton for valuable comments and advice.

Received 5 January; accepted 15 February 1983.

1. Abercrombie, M., Heaysman, J. E. M. & Pegrum, S. P. *Exp Cell Res.* **73**, 536-539 (1972).
2. Edidin, M. & Weiss, A. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2456-2459 (1972).
3. Vasiliev, J. M., Gelfand, I. M., Domnina, L. V., Dorfman, N. A. & Pletyushkina, O. Y. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4085-4089 (1976).
4. Badley, R. A., Woods, A. & Rees, D. A. *J. Cell Sci.* **47**, 349-364 (1981).
5. Abercrombie, M., Heaysman, J. E. M. & Pegrum, S. P. *Exp Cell Res.* **67**, 359-367 (1970).
6. Harris, A. K. *CIBA Fdn Symp.* **14** (new ser.), 3-26 (1973).
7. Harris, A. K. *Nature* **263**, 781-783 (1976).
8. Bretscher, M. *Nature* **260**, 21-23 (1976).
9. Hewitt, J. A. *J. theor. Biol.* **80**, 115-127 (1979).
10. Oliver, J. M. & Berlin, R. D. *Int. Rev. Cytol.* **74**, 55-94 (1982).
11. De Petris, S. & Raff, M. C. *CIBA Fdn Symp.* **14** (new ser.), 27-52 (1973).
12. Edelman, G. M. *Science* **192**, 218-226 (1976).
13. Koch, G. L. E. in *Cell Adhesion and Motility* (eds Curtis, A. S. G. & Pitts, J. D.) 425-444 (Cambridge University Press, 1980).

14. Flanagan, J. & Koch, G. L. E. *Nature* **273**, 278-281 (1978).
15. Ash, J. F. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4575-4579 (1976).
16. Ash, J. F., Louvard, D. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5584-5588 (1977).
17. Toh, B. H. & Hard, G. C. *Nature* **269**, 695-697 (1977).
18. Geiger, B. & Singer, S. J. *Cell* **16**, 213-222 (1979).
19. Heath, J. P. & Dunn, G. A. *J. Cell Sci.* **29**, 197-212 (1978).
20. Heath, J. P. *Cell Biol. int. Rep.* **5**, 975-980 (1981).
21. Heath, J. P. *J. Cell Sci.* **60** (1983).
22. Sorzano, T. & Bell, E. *J. Cell Biol.* **95**, 127-136 (1982).
23. Middleton, C. A. *Nature* **282**, 203-205 (1979).
24. Abercrombie, M., Heaysman, J. E. M. & Pegrum, S. P. *Exp Cell Res.* **60**, 437-444 (1970).
25. Harris, A. K. & Dunn, G. A. *Exp Cell Res.* **73**, 519-523 (1972).
26. Sundqvist, K.-G. & Ehrnst, A. *Nature* **264**, 226-231 (1976).
27. Albertini, D. F. & Anderson, E. J. *Cell Biol.* **73**, 111-127 (1977).
28. Bourguignon, L. Y. W. & Singer, S. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5037-5035 (1977).
29. Condeelis, J. S. *J. Cell Biol.* **80**, 751-758 (1979).
30. Abercrombie, M. *Proc. R. Soc. B207*, 129-147 (1980).
31. Dunn, G. A. in *Cell Adhesion and Motility* (eds Curtis, A. S. G. & Pitts, J. D.) 409-423 (Cambridge University Press, 1980).

Class I-like HLA genes map telomeric to the *HLA-A2* locus in human cells

Harry T. Orr

Department of Laboratory Medicine/Pathology and
Department of Microbiology, Immunobiology Research Center,
University of Minnesota 55455, USA

Robert DeMars

Laboratory of Genetics and Department of Human Oncology,
University of Wisconsin, Madison, Wisconsin 53706, USA

Among the proteins encoded by the major histocompatibility complex (MHC) are the highly polymorphic class I major transplantation antigens, HLA-A and HLA-B in man and H-2K and H-2D in mouse. Class I loci also include the less polymorphic HLA-C in man and H-2L in mouse. Class I antigens are 45,000-molecular weight (*M*_r) glycosylated membrane proteins associated on the cell surface with apparently nonpolymorphic β_2 -microglobulin¹. Linked to the murine MHC are genes encoding the class I-like Qa and Tla antigens, which are closely related structurally to H-2K, D and L, as they are also 45,000-*M*_r cell surface glycoproteins associated with β_2 -microglobulin²⁻⁷. Such class I molecules have also been found on the surface of human T lymphocytes⁸⁻¹¹ and, in one case, shown to be linked to HLA-A¹². Recombinant DNA techniques have been used to map DNA fragments to the *Tla* region in mouse^{13,14}. A similar genetic analysis of the HLA complex is hampered by the lack of congenic strains and by the small number of intra-MHC recombinants in well studied families. One means of overcoming these problems involves a molecular genetic analysis of human lymphoblastoid cell line (LCL) mutants¹⁵ having γ -ray-induced physical deletions of *HLA* DNA to map DNA fragments in the MHC¹⁶. Here we have applied this method to provide evidence that some human class I-like DNA sequences are telomeric to the *A2* locus in LCL 721.

We studied DNA from human lymphoblastoid cell line LCL 721 and mutants of it, 721.45.1 and 721.127 (unpublished), from which one entire haplotype or the other had been physically deleted by γ -irradiation. Most important, however, was the analysis of the mutant 721.144 having an A⁻ phenotype that resulted from γ -irradiation of a mutant (721.45.1) from which an entire haplotype had already been physically deleted followed by immunoselection for mutants that no longer expressed the remaining *HLA-A* allele. Figure 1 schematically describes the production of the mutants.

The restriction enzyme *Hind*III was used to digest genomic DNA of LCL 721. Southern blots of these digests were then probed with three HLA class I cDNA clones: (1) the almost full-length (amino acid 39 to the poly(A)tail) B7 cDNA clone isolated by Sood *et al.*¹⁷; (2) a *Pvu*II subclone of this clone containing the 5' 600 base pairs (bp) of coding sequence (amino acids 39-239); and (3) pHLA-1 (ref. 18), a 500-bp cDNA clone

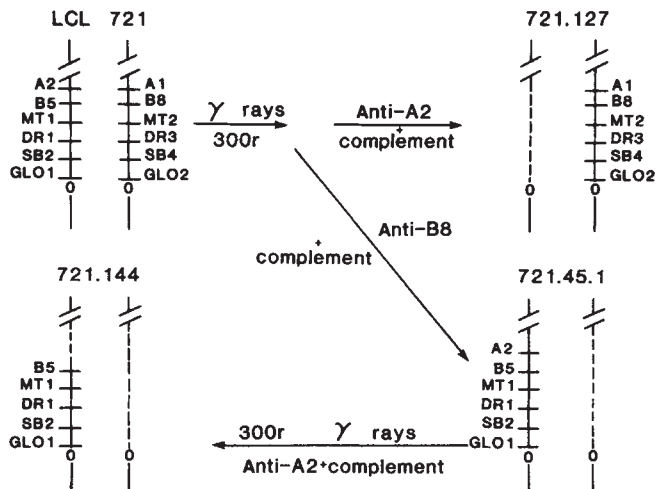


Fig. 1 The origins and phenotypes of cell mutants used. Mutants 721.45.1 and 721.127 have large microscopically visible deletions on the short arm of chromosome 6. --- Indicates the loss of phenotypic expression.

from the 3' end of a HLA class I mRNA. The 5'-specific *PvuII* subclone consists entirely of coding sequence while the 3'-specific clone contains coding sequence for only the carboxy-terminal 46 amino acids plus a portion of the 3' untranslated region. This 3' clone contains neither the poly(A) tail nor the poly(A) addition site.

The hybridization patterns obtained from the *HindIII* digest of LCL 721 DNA using these HLA cDNA probes (Fig. 2) show that the enzyme cuts relatively infrequently within HLA class I genes; 9 of 15 hybridizing bands detected with the almost full-length probe (Fig. 2b) hybridized with both the 3'-(Fig. 2a) and 5'-(Fig. 2c)-specific probes indicating that these fragments may contain complete HLA class I genes or pseudogenes.

The complementary haplotype loss mutants 721.45.1 and 721.127 and the A⁻ mutant 721.144 proved particularly useful in mapping DNA fragments to the telomeric side of A2. As expected, mutant 721.144 lacks the 4.7-kilobase(kb) *HindIII* class I hybridizing band that was detected by the B7 cDNA clone¹⁷ and found to be absent from its parental 721.45.1 cells; in addition, 721.144 lacks class I hybridizing *HindIII*-generated bands at 15.0, 7.0, 5.9, 5.1 and 2.1 kb (Fig. 3). Moreover, in mutant 721.144, the 9.2-kb, 8.0-kb and 5.5-kb hybridizing bands are clearly reduced in intensity relative to 721 and 721.45.1. The presence of the hybridizing band at 5.1 kb correlates closely with expression of A2, that is, it is specifically absent from 5 of 11 heterozygous γ ray-induced A2 loss mutants analysed so far (unpublished data). Thus, mutant 721.144 has lost class I sequences linked to A2. Localization of these class I sequences relative to A2 can be attempted by analysing the hybridization pattern obtained for mutant 721.144 with reference to those obtained for mutants 721.45.1 and 721.127. In each of the latter, DNA has been deleted from a point between the centromere and *GLO* to a point in or distal to the A locus. Therefore, the hybridization pattern for each mutant displays the class I sequences present in one or the other haplotype present in LCL 721. Except for the differences with regard to the 4.7- and 5.1-kb fragments, which result from an existing polymorphism in a cut site for *HindIII*, all fragment size classes are, at least, approximately equally represented in both haplotypes present in LCL 721. Southern blot bands absent in the A⁻ mutant 721.144 must correspond to a region of homozygous deletion that has one (proximal) boundary between the *HLA-A* and *-B* loci and the other (distal) in or distal to the *HLA-A* locus. As the locations of the distal breakpoints in 721.45.1 and 721.144 are unknown, class I sequences that are absent from 721.144 (15.0, 7.0, 5.9 and 2.1 kb) could be proximal or distal to the A locus. However, the reduced intensity of the 9.2-, 8.0- and 5.5-kb bands in 721.144 must result from (1) the deletion of these sequences

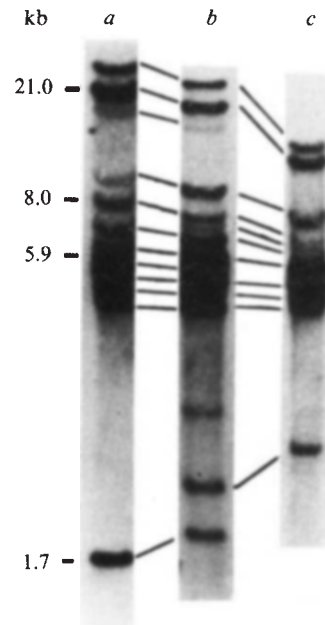


Fig. 2 Southern blot of *HindIII*-digested LCL 721 DNA using 5'- and 3'-specific HLA class I probes. a, DNA probed with the 3'-specific probe pHLA-1 (ref. 18). b, DNA probed with the nearly full-length class I cDNA clone¹⁷. c, DNA probed with a 5' 600-bp *PvuII* fragment of the Sood *et al.*¹⁷ cDNA clone containing only coding sequences (amino acids 39-239).

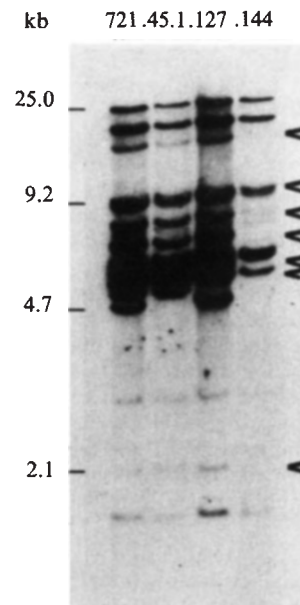


Fig. 3 Loss of HLA class I hybridizing fragments linked to A2. DNA was digested with *HindIII* and probed with the nearly full-length class I cDNA clone¹⁷. 721, DNA from LCL 721; .45.1, DNA from a mutant that has lost the *GLO2*, *DR3*, *B8*, *A1* haplotypes; .127, DNA from a mutant that has lost the *GLO1*, *DR1*, *B5*, *A2* haplotype; .144, DNA from an A⁻ mutant that was selected from mutant .45.1 after a second exposure to γ -irradiation. < Indicates bands reduced in intensity or lost in .144 but present in 721 and .45.1.

together with the A2 locus when the second deletion was introduced into 721.45.1 and (2) the retention of these sequences on the chromosome from which the A1-B8-containing haplotype had been deleted in 721.45.1. Sequences between the centromere and the proximal breakpoint of the second deletion would be in a region of homozygous deletion and be completely absent from 721.144. This means that the sequences (9.2, 8.0 and 5.5 kb) having reduced intensity in 721.144 must

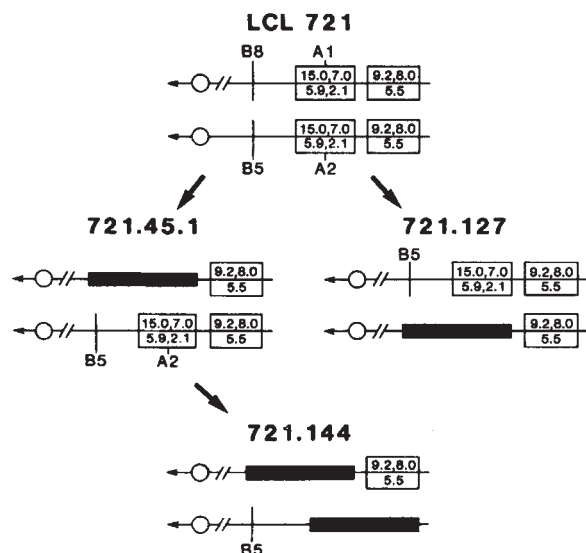


Fig. 4 Localization of HLA class I-like DNA sequences with respect to the *HLA-A* locus. This is a schematic interpretation of the data presented in Fig. 3. In addition to the fragments shown in this figure, the 4.7- and 5.1-kb fragments present in LCL 721 have been shown to be specifically associated with the A1 and A2 alleles, respectively (they have been omitted here for clarity). Solid bars represent the absence of antigens and class I DNA sequences. ○ Represents the centromere of chromosome 6. See Fig. 1 and text for the origins of the mutants.

be distal to the *A* locus (Fig. 4). We will refer to these sequences as telomerad.

In the mouse, the *Tla* region is ~1.0 cM from H-2D⁷. *Tla* (ref. 19) and *Qa-2* (ref. 20) antigens share 15–35% and 21–43%, respectively, of their tryptic peptides with H-2 antigens. These levels of peptide homology indicate that *Tla* and *Qa-2* are related to the H-2 antigens, although it is difficult to assess the exact amount of nucleotide sequence homology from such data. The presence of homology between the bands that have reduced intensity in 721.144 and the B7 cDNA clone¹⁷ as well as their location on the telomeric side of A2 suggest that at least some of the bands that are reduced in 721.144 may correspond to the murine *Tla* and/or *Qa* genes. It is possible that some of the bands specifically lost in 721.144 may also be telomeric to A2 and thus correspond to *Tla* and/or *Qa*. Presumably, the cloning of the telomerad class I DNA will provide access to an interesting group of HLA class I-like genes.

It is unknown how far beyond A2 the telomerad class I DNA is localized. However, the observation that of eight γ ray-induced mutants (unpublished data) that lost the 5.1-kb band correlating with A2, only one mutant, 721.144, also lost the 15.0-, 7.0-, 5.9- and 2.1-kb bands and had reduced intensities of the 9.2-, 8.0- and 5.5-kb bands, may imply that these class I DNA sequences are not closely linked to A2. Regardless of their distance from A2, the mapping of the 9.2-, 8.0- and 5.5-kb bands to the telomeric side of A2 complements the extension of the known boundary of the MHC towards the centromere by means of γ ray-induced segregations between *HLA-DR* and *HLA-SB*^{21,22}. Therefore, the portion of chromosome 6 in LCL 721 that contains histocompatibility genes, immune response genes and structurally related genes, may be much larger than is generally believed. Whether or not the size of the MHC as determined with the A2-containing haplotype in LCL 721 will be the same in other HLA haplotypes can be determined by dissecting further the B8-containing haplotype of LCL 721 and by applying the approaches used here to cells with other haplotypes.

This is paper 288 from the Immunobiology Research Center, University of Minnesota, and no. 2592 from the Laboratory of Genetics, University of Wisconsin, Madison. This work was supported in part by NIH grants AI-15486, P30 CA 14520 and AI-18124 and by grants from the Searle Foundation and the

Leukemia Society to H.T.O. The technical assistance of Ms C. Chang, Mr R. Rudersdorf and Mr G. Johnson is acknowledged.

Received 22 November 1982; accepted 14 January 1983.

1. Ploegh, H., Orr, H. T. & Strominger, J. L. *Cell* **24**, 287–299 (1981).
2. Rask, L., Lindblom, J. B. & Peterson, P. A. *Nature* **249**, 833–834 (1974).
3. Silver, J. & Hood, L. *Nature* **249**, 764–765 (1974).
4. Michaelson, J., Flaherty, L., Vitetta, E. & Poulik, M. *J. exp. Med.* **145**, 1066–1070 (1977).
5. Stanton, T. H. & Hood, L. *Immunogenetics* **11**, 309–318 (1980).
6. Michaelson, J. *Immunogenetics* **13**, 167–171 (1981).
7. Flaherty, L. in *The Role of the Major Histocompatibility Complex in Immunobiology* (ed. Dorf, M. E.) 33–58 (Garland STPM, New York, 1980).
8. Tada, N., Tanigaki, N. & Pressman, D. *J. Immun.* **120**, 513–519 (1978).
9. Ziegler, A. & Milstein, C. *Nature* **279**, 243–244 (1979).
10. Brodsky, F. M., Bodmer, W. F. & Parham, P. *Eur. J. Immun.* **9**, 536–545 (1979).
11. Cotner, T., Mashimo, H., Kung, P. C., Goldstein, G. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3858–3862 (1981).
12. Gazit, E., Terhorst, C. & Yunis, E. J. *Nature* **284**, 275–277 (1980).
13. Margulies, D. H., Evans, G. A., Flaherty, L. & Seidman, J. G. *Nature* **295**, 168–170 (1982).
14. Pease, L. R., Nathenson, S. G. & Leinwand, L. A. *Nature* **298**, 382–385 (1982).
15. Kavathas, P., Bach, F. H. & DeMars, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4251–4255 (1980).
16. Orr, H. T. *et al. Nature* **296**, 454–456 (1982).
17. Sood, A. K., Pereira, D. & Weissman, S. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 616–620 (1981).
18. Ploegh, H. L., Orr, H. T. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6081–6085 (1980).
19. Yokoyama, K., Stockert, E., Old, L. V. & Nathenson, S. G. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7078–7082 (1981).
20. Soloski, M. J., Uhr, J. W. & Vitetta, E. S. *Nature* **296**, 759–761 (1982).
21. Shaw, J., Kavathas, P., Pollack, M. S., Charmot, D. & Mawas, C. *Nature* **293**, 745–747 (1981).
22. Kavathas, P., DeMars, R., Bach, F. H. & Shaw, S. *Nature* **293**, 747–749 (1981).

Cloning and expression in *E. coli* of the malarial sporozoite surface antigen gene from *Plasmodium knowlesi*

Joan Ellis, Luiz S. Ozaki, Robert W. Gwadz, Alan H. Cochrane, Victor Nussenzweig, Ruth S. Nussenzweig & G. Nigel Godson

The Department of Biochemistry, The Division of Parasitology of The Department of Microbiology and the Department of Pathology, New York University Medical Center, 550 First Avenue, New York, New York 10016, USA and The Laboratory of Parasitic Diseases, National Institutes of Health, Bethesda, Maryland 20205, USA

The malarial sporozoite, the infective stage found in the salivary gland of the insect vector, bears highly immunogenic surface antigen(s). Repeated exposure to irradiated sporozoites induces protection against malaria in several host species, including man¹. Further, monoclonal antibodies that confer passive immunity react with the immunogenic surface determinants of different sporozoite species^{2–4}. One approach to prevent malaria, therefore, would be to produce a vaccine that induces high titres of circulating antibodies against the sporozoite surface determinant(s). However, production of such a vaccine has not been possible since sporozoites cannot be cultivated *in vitro* and, therefore, only limited amounts of surface antigen may be obtained. To overcome this problem, we have prepared mRNA from *Plasmodium knowlesi*-infected mosquitoes to construct a cDNA library. From this library we have isolated a clone that expresses the sporozoite surface antigen as a β -lactamase fusion protein in the plasmid pBR322. This is the first potentially protective malarial antigen to be cloned by recombinant DNA technology.

To demonstrate the feasibility of obtaining a cDNA clone to the sporozoite surface antigen, poly(A)⁺ mRNA from *P. knowlesi*-infected thoraxes of *Anopheles dirus* mosquitoes was purified by oligo d(T) cellulose chromatography⁵, and translated *in vitro* in a wheat germ cell-free system⁶. The products of the *in vitro* translation were immunoprecipitated with 2G3, a monoclonal antibody specific for the protective surface antigen on the *P. knowlesi* sporozoite³ (Fig. 1A). A polypeptide of $M_r = 51,000$ was immunoprecipitated from the *in vitro* translation products of mRNA derived from infected mosquitoes (Fig. 1A, lane 3), but not from the translation products of mRNA from

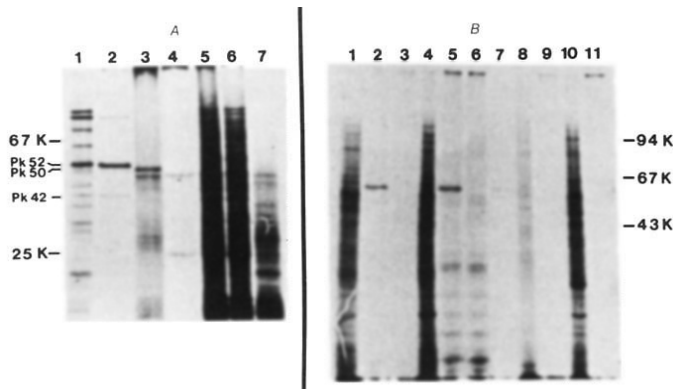


Fig. 1 A, Lane 1, total *in vivo* ^{35}S -methionine labelled sporozoites. Lane 2, same as lane 1, but immunoprecipitated with 2G3. Lane 3, total translated products from *P. knowlesi*-infected thoraxes labelled with ^{35}S -methionine and immunoprecipitated with 2G3. Lane 4, total translated polypeptides from uninfected mosquito thoraxes labelled with ^{35}S -methionine and immunoprecipitated with 2G3. Lane 5, total translated polypeptides from *P. knowlesi*-infected thoraxes labelled with ^{35}S -methionine. Lane 6, total translated polypeptides from uninfected mosquito thoraxes. Lane 7, total wheat germ extract. B, Lane 1, total polypeptides translated from RNA of *P. knowlesi*-infected thoraxes labelled with ^{35}S -methionine. Lane 2, same as lane 1, but immunoprecipitated with 2G3. Lane 3, same as lane 1, but immunoprecipitated with 3D11. Lane 4, total polypeptides translated from mRNA of *P. knowlesi*-infected thoraxes labelled with ^{35}S -methionine. Lane 5, same as lane 4, but immunoprecipitated with 2G3. Lane 6, same as lane 4, but immunoprecipitated with 3D11. Lane 7, translated polypeptides from RNA of column-purified sporozoites labelled with ^{35}S -methionine and immunoprecipitated with 2G3. Lane 8, total translated polypeptides from *P. knowlesi* blood stage RNA. Lane 9, same as lane 8, but immunoprecipitated with 2G3. Lane 10, total polypeptides translated from RNA of uninfected thoraxes labelled with ^{35}S -methionine. Lane 11, same as lane 10, but immunoprecipitated with 2G3. **Methods:** Immunoprecipitation of translated and *in vivo* ^{35}S -methionine labelled-*P. knowlesi* polypeptides. RNA was extracted from infected and uninfected *Anopheles dirus*^{15,16}. Both total RNA and oligo d(T) cellulose-purified mRNA were translated in a wheat germ system⁶. Translated products were labelled in the presence of ^{35}S -methionine and immunoprecipitated with either the sera of mice bearing the *P. knowlesi* hybridoma, 2G3, or the *P. berghei* hybridoma, 3D11^{18,19}. The precipitated products were separated by SDS-polyacrylamide gel electrophoresis, and the dried gel autoradiographed²⁰. *In vivo* sporozoite extracts were prepared from *Anopheles dirus* salivary glands dissected in M199 with 1% bovine serum albumin. Sporozoites were purified over lectin columns and labelled in RPMI-1640 in the presence of ^{35}S -methionine for 2 h at room temperature²¹. The reaction was stopped by the addition of 10 mM methionine, 0.05% sodium azide and 2 mg ml⁻¹ BSA. The sporozoites were then lysed and extracted with 1% Nonidet P-40 for immunoprecipitation. The labelled proteins were then analysed by SDS-polyacrylamide gel electrophoresis, and autoradiography.

uninfected mosquitoes (Fig. 1A, lane 4). The *in vitro*-synthesized polypeptide is, therefore, related antigenically to the three proteins ($M_r = 52,000$, 50,000 and 42,000) that are immunoprecipitated from *P. knowlesi* sporozoites labelled *in vivo* with ^{35}S -methionine (Fig. 1A, lane 2). Previous studies have shown that the $M_r = 42,000$ polypeptide is represented on the surface of sporozoites while the higher molecular weight products are intracellular precursors³. The immunoprecipitation of a single *in vitro* product by the same monoclonal antibody that recognizes three *in vivo*-labelled polypeptides suggests that a single mRNA molecule encodes all these polypeptides. This is supported by the finding that a single mRNA band is hybridized on the Northern gel of infected mosquito mRNA when cloned cDNA was used as a probe (see below and Fig. 2). The difference

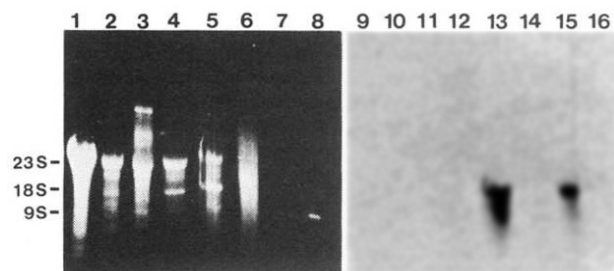


Fig. 2 Northern blot analysis of pEG81. mRNAs from uninfected and *P. knowlesi*-infected mosquito thoraxes and from *P. knowlesi* blood forms were prepared by standard methods^{15,16}. mRNA from *P. knowlesi*-infected mosquito thoraxes was also prepared by a polysome precipitation procedure¹⁴ with anti-*P. knowlesi* monoclonal antibody, 2G3. Lanes 1-8 show the ethidium bromide-stained RNA patterns and lanes 9-16, the autoradiograph of the filter. Lanes 1 and 9, 16 μg *Anopheles dirus* rRNA. Lanes 2 and 10, 16 μg infected *A. dirus* rRNA. Lanes 3 and 11, 10 μg rabbit reticulocyte rRNA. Lanes 4 and 12, 5 μg uninfected mosquito thorax mRNA. Lanes 5 and 13, 5 μg infected mosquito thorax mRNA. Lanes 6 and 14, 5 μg *P. knowlesi* blood stage mRNA. Lanes 7 and 15, polysome precipitated mRNA from *P. knowlesi*-infected mosquito thoraxes mixed with globin mRNA (BRL) as a carrier. Lanes 8 and 16, 0.15 μg globin mRNA. **Methods:** The mRNA samples were separated on a 10 mM methyl mercury - 1.2% agarose slab gel. The gel was washed twice in 0.2 M sodium acetate (pH = 4.0) and blotted overnight onto activated ABM paper (Scheicher and Schell). The filter was pre-washed for 2 h at 65 °C in 3 $\times$ SSC, 10 $\times$ Denhardt's¹⁷, 1% glycine, 0.1% SDS, and hybridized overnight at 65 °C with the probe ^{32}P -labelled pEG81. The hybridized filter was extensively washed in SSC in gradually decreasing concentrations from 2 $\times$ to 0.1 $\times$ at 65 °C.

in electrophoretic mobility between the *in vitro* synthesized product and the polypeptides detected after *in vivo* labelling is probably due to post-translational modifications of the $M_r = 51,000$ polypeptide⁷.

The $M_r = 51,000$ polypeptide showed the species and stage specificity characteristic of the sporozoite surface antigen⁸. The translation products of mRNA derived from infected thoraxes were immunoprecipitated by the *P. knowlesi* monoclonal antibody, 2G3, (Fig. 1B, lanes 2, 5), but not by the *P. berghei* monoclonal antibody, 3D11 (Fig. 1b, lanes 3 and 6). Our data also support the concept of stage specificity of the sporozoite antigen since we were unable to immunoprecipitate a polypeptide with 2G3 from translation systems programmed with blood stage mRNA (Fig. 1B, lane 9) and we were also unable to identify a blood stage RNA product on a Northern gel when the cloned cDNA is used as a probe (see below and Fig. 2).

To isolate the sporozoite surface antigen gene, a cDNA library was prepared from mRNA of *P. knowlesi*-infected mosquito thoraxes. mRNA was converted to cDNA with reverse transcriptase⁹, size-fractionated on a Sepharose 4B column, tailed with deoxycytidine, annealed to deoxyguanine-tailed plasmid pBR322¹⁰ and used to transform *E. coli* RR1 cells¹¹. Tetracycline-resistant *E. coli* colonies were screened for the production of *P. knowlesi* antigens with monoclonal antibodies in a two-site immunoradiometric assay.¹² Crude cell extracts¹³ of 48-colony pools were absorbed onto microtitre plates pre-coated with a pool of three different surface antigen specific monoclonal antibodies, 6B8, 8B8 and 8E11.³ The plates were washed and bound antigen was revealed with ^{125}I -labelled surface antigen-specific monoclonal antibodies 2G3 and 8A8³. Colonies from positive pools were individually screened for antigen production and lysates from three clones pEG81, pEG117, and pEG118, were found to bind the ^{125}I -labelled antibodies. The sizes of the cDNA inserts were approximately 350, 1,200 and 1,200 base pairs respectively. Southern blot analysis has shown that the three clones cross-hybridize and initial DNA sequencing results indicate that pEG117 and pEG118 are identical.

The specificity of the proteins encoded by these three clones

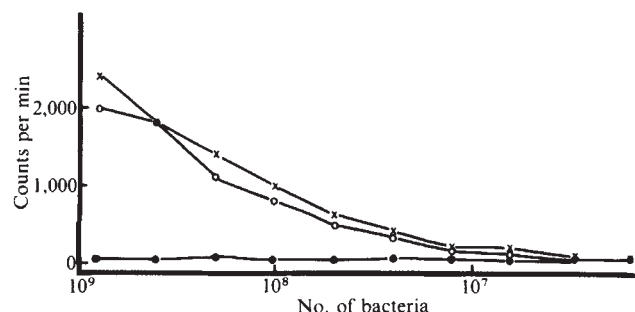


Fig. 3 Characterization of pEG117 by competition-inhibition assay. Cultures of pEG117-containing cells were lysed in 0.1 M Tris, pH 8.0 and 5 mg ml⁻¹ lysozyme on ice for 30 min. The extracts were sonicated briefly to ensure total cell lysis. To perform the immunoradiometric assay microtitre wells were coated with 50 µl (50 µg ml⁻¹ in phosphate-buffered saline) anti-*P. knowlesi* monoclonal antibody, 2G3, overnight at 4°C. After repeated washing with PBS-1% BSA, the wells were incubated with two-fold serial dilutions of pEG117 extract for 4 h at room temperature. The wells were thoroughly washed again and incubated with 10 µg of unlabelled 2G3 in 25 ml of PBS-BSA (●) or with 10 µg of unlabelled 3D11 in 25 µl of PBS-BSA (○) or with 25 µl of PBS-BSA only (×) for 1 h. 1 ng of ¹²⁵I-labelled 2G3 in 25 µl of PBS-BSA was added to all the wells, the plates were incubated for another 2 h at room temperature, washed and the wells counted.

was tested further by competition-inhibition immune assays. Crude cell extracts of *E. coli* containing pEG81, pEG117 or pEG118 were assayed as above in the presence or absence of unlabelled antibody (Fig. 3). All *P. knowlesi* antibody binding was blocked by pretreatment with unlabelled 2G3, but this binding was unaffected by pre-treatment with unlabelled 3D11, an unrelated monoclonal antibody.

In these experiments the same monoclonal antibody could be used on top and on the bottom in the sandwich immune assay, suggesting that the antigen contains at least two epitopes recognized by a single monoclonal antibody. Consistent with this observation, preliminary DNA sequence data suggest the occurrence of a repeated peptide segment within the surface antigen protein.

To prove that immunoreactive clones were of parasite and not of mosquito origin, the purified ³²P-labelled pEG81 plasmid DNA was used as a probe against uninfected mosquito and blood stage DNA in a Southern blot analysis. The cloned DNA hybridized with an *Eco*RI fragment 12–14 kb long from blood stage DNA, but did not hybridize with uninfected mosquito or *E. coli* DNA (data not shown). The ³²P-labelled pEG81 DNA was also used to probe a Northern blot of mRNA from blood stage and infected and uninfected mosquitoes. The cloned DNA hybridized with a single band of approximately 2,050 nucleotides present in infected mosquito mRNA and in mRNA from immunoprecipitated sporozoite polysomes¹⁴ (Fig. 2). This size difference between the mRNA and the cDNA insert indicates that none of the three clones contains the full length gene. However, since the monoclonal antibody is able to react with the crude cell extracts, these clones must encode and express the antigenic determinant of the sporozoite surface protein. As expected, uninfected mosquito mRNA and blood stage mRNA did not hybridize with the probe.

With the immunoreactive part of the sporozoite surface antigen gene cloned and expressed in *E. coli* as a β-lactamase fusion protein we now should have sufficient antigen for both further structural analysis, and the development of a model malaria vaccine directed against the sporozoite stage which should be helpful in developing a human vaccine.

We thank Dr David Colman for helping with the RNA extraction and the *in vitro* translation techniques; Drs Milton Adesnik, Warren Jelinek and Fidel Zavala for their helpful suggestions, and Drs Peter D'Eustachio and David Colman for critical reading of the manuscript. We also thank P. Svec, C. Raphael, R. Altzuler and M. Marancic for technical assistance.

This investigation was supported by the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (T16/181/M2/21(G)) (G.N.G.), New York University Institutional Funds (G.N.G.), The Rockefeller Foundation (R.S.N.), Agency for International Development (R.S.N.) and by a Training Grant from the NIH (J.E.).

Received 24 January; accepted 23 February 1983.

1. Nussenzweig, R. S. in *Immunity to Blood Parasites of Animals and Man* (eds Miller, L., Pino, J. & McKelvey, J.) 75–87 (Plenum, New York, 1977).
2. Yoshida, N., Nussenzweig, R. S., Potocnjak, P., Nussenzweig, V. & Aikawa, M. O. *Science* **207**, 71–73 (1980).
3. Cochran, A., Santoro, F., Nussenzweig, V., Gwazd, R. & Nussenzweig, R. S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5651–5655 (1982).
4. Nardin, E. H. *et al. J. exp. Med.* **156**, 20–30 (1982).
5. Aviv, H., & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
6. Roman, R., Brooker, J. D., Seal, S. N. & Marcus, A. *Nature* **260**, 359–360 (1978).
7. Steiner, D. F. *et al. Prog. Cell Biol.* **4**, 175–202 (1980).
8. Cochran, A. H., Nussenzweig, R. S. & Nardin, E. H. in *Malaria* Vol. 3 (ed. Kreier, J.) 163–202 (Academic, New York, 1980).
9. Efstratiadis, A., Kafatos, F. C., Maxam, A. M. & Maniatis, T. *Cell* **7**, 279–288 (1976).
10. Villa-Komaroff, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 3727–3721 (1978).
11. Dagert, M. & Erlich, S. D. *Gene* **6**, 23–28 (1979).
12. Zavala, F., Gwazd, R. S., Collins, F. H., Nussenzweig, R. S. & Nussenzweig, V. *Nature* **299**, 737–738 (1982).
13. Carbon, J., Clarke, L., Chinault, C., Ratzkin, B. & Walz, A. in *Biochemistry and Genetics of Yeasts* (eds Bacila, M., Horecker, B. L. & Styspani, A. O. M.) 425–443 (Academic, New York, 1978).
14. Korman, A. J., Knudsen, P. J., Kaufman, J. F. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1844–1848 (1982).
15. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
16. Liu, C., Slate, D. L., Gravel, R. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4503–4506 (1979).
17. Denhardt, D. *Biochem. biophys. Res. Commun.* **23**, 641–652 (1966).
18. Roberts, B. E. & Paterson, B. M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2330–2334 (1973).
19. Goldman, B. M. & Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5066–5070 (1978).
20. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
21. Yoshida, N., Totocnjak, P., Nussenzweig, V. & Nussenzweig, R. S. *J. exp. Med.* **154**, 1225–1236 (1981).

Isolation and nucleotide sequence of a cDNA encoding the precursor of mouse nerve growth factor

James Scott*†, Mark Selby*, Mickey Urdea‡, Margarita Quiroga‡, Graeme I. Bell‡ & William J. Rutter*

* Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA
‡ Chiron Corporation, 4560 Horton Street, Emeryville, California 94608, USA

Nerve growth factor (NGF) is a polypeptide that enhances survival, nerve fibre outgrowth and neurotransmitter biosynthesis in sympathetic and sensory neurones^{1–3}. Administration of antibodies against NGF to developing animals leads to atrophy of the sympathetic system⁴. NGF is not normally detectable in innervated tissues but ablation of the innervating neurones leads to the production of measurable NGF in the target tissue⁵. After transplantation of the denervated tissue, reinnervation occurs, then NGF decreases to undetectable levels. Thus NGF seems to act as a neurotrophic messenger and its level is regulated by innervating neurones. Because of the minute levels present it is very difficult to study NGF biosynthesis in innervated tissue. However, NGF can be isolated from male mouse submaxillary glands, where it exists in inexplicably high levels⁶. Its amino acid sequence has been determined⁷, and the synthesis of NGF and its larger precursors has been demonstrated in cultured submaxillary glands⁸. We report here the nucleotide sequence of a submaxillary cDNA encoding the mouse NGF precursor (preproNGF). In contrast to previous suppositions⁹ the NGF moiety is situated near the carboxy-terminus of the polypeptide precursor. It is flanked at the amino-terminus by 187 amino acids which may be cleaved at dibasic residues to generate three peptides; there are only two additional amino acids at the carboxy-terminus.

† Present address: Molecular Medicine Group, MRC, Clinical Research Centre, Harrow, Middlesex HA1 3UJ, UK.

GAGCGCCUGGAGCCGAGGGGAGCGCAUCGAGACUUGGAGCUGGCCUUUAUUUUGG 59

AUCUCCCGGGCAGCUUUUUGGAAACUCCUAGUGAAC 110

10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 307

pro val lys leu gly ser leu glu val gly his gly gln cau gly 155

GGA GUU UUG GCC UGU GGU CGU GCA GUC gln GGG GCU GGA UGG CAU 200

ala gly pro lys leu thr ser val ser gly pro asn lys gly phe 245

GCU GGA CCC AAG CUC ACC UCA GUG UCU GGG CCC AAU AAA GGU UUU

ala lys asp ala glu phe tyr thr gly arg ser glu val his ser 290

GCC AAG GAC GCA GAA UUC UAU ACU GGC CGC AGU GAG GUG CAU AGC

val met ser met leu phe tyr thr leu ile thr ala phe leu ile 335

GUA AUG UCC AUG UUG UUC UAC ACU CUG ACU GCG UUU UUG AUC

gly val gln ala glu pro tyr thr asp ser asn val pro glu gly 380

GCG CAG GAG GCA GAA CCG GAA GAC ACA GAG GAG GAG GAG GAG GAG

asp ser val pro glu ala his trp thr lys leu cau ser leu 425

GAC UCU GUC CCU GAA GCC CAC UGG ACU AAA CUU CAG CAU UCC CUU

asp thr ala leu arg arg ala arg ser ala pro thr ala pro ile 470

GAC ACA GCC CUC CGC AGA GCC CGC AGU GGC CCU ACU GCA CCA AUA

ala ala arg val thr gly gln thr arg asn ile thr val asp pro 515

GCU GCC CGA GUG ACA GGG CAG ACC CGC AAC AUC ACU GUA GAC CCC

arg leu phe lys lys arg arg leu his ser pro arg val leu phe 560

AGA CUG UUU AAG AAA CGG AGA CUC CAC UCA CCC CGU GUG CUG UUC

ser thr gln pro pro thr ser ser asp thr leu asp leu asp 605

AGC ACC CAG CCU CCA CCC ACC UCU UCA GAC ACU CUG GAU CUA GAC

phe gln ala his gly thr ile pro phe 180 650

UUC CAG GCC CAU GGU ACA AUC CCU UUC AAC AGG ACU CAC CCG ser

lys arg ser ser thr his pro val phe his met gly glu phe ser 695

AAG CGC UCA UCC ACC CAC CCA GUC UUC CAC AUG GGG GAG UUC UCA

val cys asp ser val ser val trp val gly asp lys thr thr ala 740

GUG UGU GAC AGU GUG GUG GUU GGA GAG GAU AAG ACC ACA GCC

thr asp ile lys gly lys glu val thr val leu ala glu val asn 785

ACA GAC AUC AAG GGC AAG GUG GUG ACA GAG CUG GAG GAG val AAC

ile asn asn ser val phe arg gln tyr phe phe glu thr acc lys cys 830

AUU AAC AAC AGU GUA UUC AGA CAG UAC UUU UUU UUU GAG ACC AAG UGC

arg ala ser asn pro val glu ser gly cys arg gly ile asp ser 875

CGA GCC UCC AAU CCU GUU GAG AGU GGG UGC CGG GGC AUC GAC UCC

lys his trp asn ser tyr cys thr thr thr his thr phe val lys 920

AAA CAC UGG AAC UCA UAC UGC ACC ACU ACU CAC ACC UUC UUC AAG

ala leu thr thr asp glu lys gln ala ala trp arg phe ile arg 965

GCG UUG ACA ACA GAU GAG AAG CAG GCU GCC UGG AGG UUC AUC CGG

ile asp thr ala cys val cys val leu ser arg lys ala thr arg 1010

AUA GAC ACA GCC UGU GUG UGU GUG CUC AGC AGG AAG GCU ACA AGA

arg gly OP 1067

AGA GGC UGA CUUGCCUGACGCCCUUCCCCACCUGCCCUCCACACUCUCUUGG

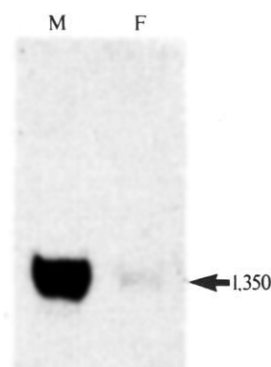
CCCCUCCUACCCAGCCUGUAAUUUUUUUUUUUUUUUAAAGACUGCAUGAUUUUUU 1,127

GUUUUACAAUUUUUAAAGACUUUUUUUUUUUUUUUUUUAAGCAUCCUGAAAAAAA

Fig. 1 Nucleotide sequence of the mouse submaxillary gland NGF mRNA and the derived amino acid sequence of the NGF precursor. The NGF mRNA sequence was determined by sequencing both strands of a 1,068-bp cDNA insert and from the sequence of cDNA generated by a primer extension reaction¹⁶. The number of nucleotides in the cDNA is indicated at the end of each line. The predicted amino acid sequence of mouse preproNGF is numbered by designating the first methionine as 1. The sequence corresponding to the mature NGF is labelled and underlined.

NGF-specific clones were isolated from a male submaxillary gland cDNA library by screening with synthetic oligodeoxynucleotide probes synthesized according to the nucleotide sequence predicted from the NGF amino acid sequence. To construct the library, poly(A)-containing RNA was isolated from the submaxillary glands of 60-day-old male Swiss-Webster mice^{9,10}. Double-stranded cDNAs were prepared and inserted

Fig. 2 Size determination of NGF mRNA from male and female mice; 10 µg of glyoxylated RNA obtained from male and female submaxillary glands were separated on a 2% agarose gel, transferred to a nitrocellulose filter and hybridized with nick-translated ³²P-labelled NGF cDNA insert²¹. After washing, the RNA blot was autoradiographed at -70 °C using an intensifying screen. Lane M, male RNA; F, female RNA. Glyoxylated *Hae*III-digested ΦX 174 fragments were used as size markers.



into the *Pst*I site of a pBR322 derivative (given by Dr R. A. Hallelwell, Chiron Corporation) using the G-C tailing technique¹⁰. Tetracycline-resistant transformants of *Escherichia coli* HB101 were stored at -70 °C in microtitre dishes; 5,000 of the transformants were grown on Whatman 541 filter paper, the plasmids were amplified *in situ* with chloramphenicol and the DNA immobilized on the filters^{11,12}. The oligonucleotide probe sequences used were constructed from the hexapeptide having the least degenerate codons within the NGF sequence: Gln-Tyr-Phe-Phe-Glu-Thr (excluding the third base of Thr). A mixture of all 32 possible 17-base oligonucleotides (3'-

GT^TCAT^AGAA^AGAA^ACT^TCTG-5') was synthesized by using

solid-phase phosphoramidite methodology followed by isolation from 20% acrylamide gels (ref. 13 and M. Urdea *et al.*, in preparation). ³²P-end-labelled probes were used to screen the cDNA library¹⁴. Six colonies yielded strong signals with this probe; one contained a plasmid with an insert of ~1,100 base pairs (bp). This insert was used to rescreen the library and revealed one additional hybridizing colony. The relative abundance of NGF clones in the library (7 in 5,000) suggests that NGF mRNA comprises ~0.1% of the mRNA from this source.

Most of the amino acids encoded by NGF mRNA were deduced from the nucleotide sequence of the longest cDNA clone, which includes 1,068 bases complementary to NGF mRNA (Fig. 1, nucleotides 109–1,176), a 50-base poly(A) tract, in addition to the poly(dC) and poly(dG) tails¹⁵. The cDNA was used as a probe to establish the size of the NGF mRNA in both male and female submaxillary glands. Figure 2 shows that it is ~1,350 bases and the relative level of NGF mRNA is about 10-fold more abundant in the male. Thus, the well known difference in the levels of NGF protein between males and females is due to regulation at the level of mRNA. The size of the NGF mRNA suggests that the cDNA is missing 100–150 bp at the 5' end. The remainder of the sequence was determined from a cDNA made by primed synthesis on the male mRNA using a ³²P-labelled 15-base oligonucleotide, 3'-CCACCTCAAAACCGG-5', which is complementary to the mRNA segment encoding nucleotides 153–167 (Fig. 1)¹⁶. The sequence of the extended primer confirmed the region in common with the cDNA and provided an additional 108 bases. This RNA has a 5' untranslated region of 100 bases, a single open reading frame which starts at the first methionine codon (nucleotides 96–98), and ends with the termination codon TGA (nucleotides 1,117–1,119). This nucleotide sequence predicts a 307-amino acid polypeptide which contains NGF (amino acids 188–305). The terminal 91 bases of the 160-base 3' untranslated region is exceptionally A+T-rich (80% A+T); the variant polyadenylation signal, AUUAAA, is found 22 nucleotides from the poly(A) tail¹⁷.

In common with other secreted proteins, NGF probably has an amino-terminal sequence which signals the translocation of the nascent protein into the cisternae of the endoplasmic reticulum (Fig. 3)¹⁸. Such signal peptides are usually 15–30 amino acids long, contain a central hydrophobic region and terminate in an amino acid possessing a small neutral side chain.



Fig. 3 Schematic representation of the mouse submaxillary preproNGF polypeptide. The basic di- and tetrapeptide possible cleavage sites are indicated; the sites at which cleavage is known to occur are indicated by solid boxes. The numbers in parentheses are tentative size designations; the others give accurate sizes.

The amino-terminal region of the NGF precursor contains two weakly hydrophobic regions (residues 7–14 and 22–23). The boundary of the signal peptide cannot be predicted from the amino acid sequence, but it will be evident when the amino-terminus of proNGF is determined. The predicted molecular weights of preproNGF ($M_r = 33,800$) and proNGF ($M_r = 30,000$ – $32,000$) are in reasonable agreement with the values of $\sim 36,000$ and $29,000$, respectively obtained for putative higher molecular weight precursors by T. Darling and E. Shooter (personal communication). The amino acid sequence predicted from the cDNA agrees with the sequence of Angeletti *et al.*⁷ except at residue 30, in which position our sequence predicts an aspartate. Angeletti *et al.*⁷ obtained equivocal evidence supporting the assignment of Asp or Asn at this residue; they chose Asn because of the possibility of deamidation. Therefore, it seems likely that this difference is due to an amino acid sequencing error, but the possibility of a DNA polymorphism involving a single base change cannot be ruled out.

Formation of NGF from its precursor requires proteolytic processing at both ends (Fig. 3). Cleavage sites are: Arg-Arg, 115–116; Lys-Lys-Arg-Arg, 144–147; Lys-Arg, 186–187 at the amino-terminus of NGF; and Arg-Arg, 305–306 at its carboxy-terminus. The residues Arg-Lys, 301–302, exist within the NGF sequence itself, but these residues are not apparently cleaved; they may be stabilized by the diacidic residues Asp-Glu at positions 280–281. Cleavage at Arg-Arg, 115–116 and Lys-Lys-Arg-Arg, 144–147 would result in NGF-containing moieties of M_r 22,000 and 18,000, respectively.

The extent and ordering of these possible processing reactions is unknown, but Berger and Shooter⁸ have found a relatively stable 22,000- M_r NGF precursor, which must represent active cleavage at residues 115–116. Cleavage at all processing sites in the amino-terminal region would yield three peptides of ~ 90 amino acids (residues 25–114), 27 amino acids (residues 117–143), and 38 amino acids (residues 148–186). A computer search of a compendium of protein sequences revealed no homology between these peptides and other known sequences¹⁹. The functional role of these cryptic peptides is unknown, but these sequences obviously cannot represent the 26,000- M_r α - and γ -subunits of the 7S complex¹. Perhaps proNGF is a member of the growing class of polyproteins that are processed into functionally distinct entities. The availability of the sequence of these peptides will allow direct examination of their biological properties.

There are several arginine-specific esterasepeptidases in the submaxillary gland which could perform the proNGF processing reactions; one of these is the γ -subunit of the 7S complex. This esterasepeptidase is bound to the carboxy-terminal arginine of the NGF monomer, and is probably the enzyme involved in processing the two carboxy-terminal amino acids. It seems likely that this esterasepeptidase cleaves the amino-terminal sites also.

As reported previously, the amino acid sequence of mature NGF is approximately 25% homologous with proinsulin²⁰. Alignment of the corresponding mRNA segments indicated that the nucleotide sequences possess 40% homology. However, the overall structure of the cDNAs and the positioning of the processing sites in the precursors are quite different. Thus, convergent evolution is as attractive as divergent evolution in explaining the structural resemblance between these molecules. The location of intervening sequences in the gene structure may help to decide this issue.

The availability of cDNA probes will facilitate the unambiguous determination of the sites of synthesis of NGF and the study of NGF gene expression during development, including

the puzzling role of androgens in regulating the expression of this gene in the submaxillary gland. In addition, now it should be possible to isolate related genetic sequences from the same or different species, and to examine the role of NGF in diseases of the human nervous system.

We thank Drs Tom Darling and Eric Shooter for helpful discussions and communication of unpublished results, Drs Charles Craik and Richard Najarian for assisting with computer searches, Noel Fong for technical assistance, and Maureen Appling and Leslie Spector for typing the manuscript. J.S. was supported in part by the EMBO. This research was supported by NIH grant 21344 (to W.J.R.).

Received 21 December 1982; accepted 27 January 1983.

1. Bradshaw, R. A. *Rev. Biochem.* **47**, 191–216 (1978).
2. Thoenes, H. & Barde, Y. A. *Physiol. Rev.* **60**, 1284–1335 (1980).
3. Yanker, B. A. & Shooter, E. M. *Rev. Biochem.* **51**, 849–868 (1982).
4. Levi-Montalcini, R. & Booker, B. *Proc. natn. Acad. Sci. U.S.A.* **46**, 384–391 (1960).
5. Ebendal, T., Olson, L., Seiger, A., Hedlund, K. O. *Nature* **286**, 25–28 (1980).
6. Cohen, S. *Proc. natn. Sci. U.S.A.* **46**, 303–311 (1960).
7. Angeletti, R. H., Hermodson, M. A. & Bradshaw, R. A. *Biochemistry* **12**, 100–115 (1973).
8. Berger, E. A. & Shooter, E. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3647–3651 (1977).
9. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
10. Goodman, H. M. & MacDonald, R. J. *Meth. Enzym.* **68**, 75–90 (1980).
11. Gergen, J. P., Stern, R. H. & Wensink, P. C. *Nucleic Acids Res.* **7**, 2115–2136 (1979).
12. Ish-Horowitz, D. & Burke, J. F. *Nucleic Acids Res.* **9**, 2989–2998 (1981).
13. Beaucage, S. L. & Caruthers, M. H. *Tetrahedron Lett.* **22**, 1859–1862 (1981).
14. Wallace, R. B. *et al. Nucleic Acids Res.* **9**, 879–894 (1981).
15. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
16. Argwal, K. L., Brunstedt, J. & Noyes, B. E. *J. biol. Chem.* **256**, 1023–1028 (1981).
17. Hagenbuchle, O., Bovey, R. & Young, R. A. *Cell* **21**, 179–187 (1980).
18. Lingappa, V. R. & Blobel, G. *Recent Prog. Horm. Res.* **36**, 451–475 (1980).
19. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Spring, Maryland, 1978).
20. Frazier, W. A., Angeletti, R. & Bradshaw, R. A. *Science* **176**, 482–488 (1972).
21. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).

Three mutant insulins in man

S. Shoelson, M. Haneda, P. Blix, A. Nanjo*, T. Sanke*, K. Inouye†, D. Steiner, A. Rubenstein & H. Tager‡

Departments of Biochemistry and Medicine, The University of Chicago, Chicago, Illinois 60637, USA

* Department of Medicine, Wakayama Medical School, Wakayama City, Japan

† Shionogi Research Laboratory, Osaka, Japan

We have previously identified a structurally abnormal insulin in the serum and pancreas of a middle-aged man with diabetes mellitus^{1,2} which arose from a leucine for phenylalanine substitution at position 24 or 25 of the insulin B chain; further analysis of the patient's leukocyte DNA showed that one of the patient's insulin alleles had undergone mutation resulting in loss of an *Mbo*II restriction site normally present in the human insulin gene³. Two additional and unrelated patients with the same clinical syndrome have now been identified (ref. 4 and unpublished results). All of these patients showed hyperglycaemia typical of diabetes and with marked hyperinsulinaemia typical of insulin resistance, but all three show normal tolerance to exogenously administered insulin. As the opportunity of examining pancreatic tissue from patients suspected of secreting insulin variants is rare, we have developed a method combining HPLC and radioimmunoassay to identify insulin variants isolated from human sera. By this method we have shown that all three patients noted above secrete structurally variant and chemically distinct insulins. In correction of our original assignment, one is identified as [Leu^{B25}]insulin.

‡ To whom correspondence should be addressed.

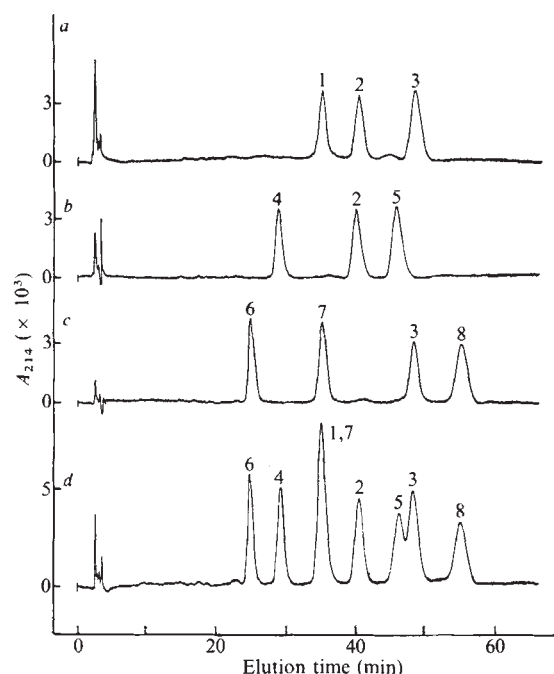


Fig. 1 Separation of insulin analogues by reverse-phase HPLC. *a*, Natural insulins differing by single amino acid residues at position B30: 1, rabbit insulin (Ser^{B30}); 2, human insulin (Thr^{B30}); 3, pig insulin (Ala^{B30}). *b*, Semisynthetic analogues of human insulin with single amino acid substitutions: 4, human [Leu^{B25}]insulin; 2, natural human insulin; 5, human [Leu^{B24}]insulin. *c*, Semisynthetic analogues of pig insulin with single amino acid substitutions: 6, pig [Ala^{B24}]insulin; 7, pig [Leu^{B25}]insulin; 3, natural pig insulin; 8, pig [Leu^{B24}]insulin. *d*, Chromatography of a mixture of all eight natural and semisynthetic insulins identified above. Rabbit and human insulins were purified from pancreatic tissue. Pig insulin was obtained from Novo Laboratories, Copenhagen and semisynthetic insulin analogues were prepared as described^{5,7,9}. The reverse-phase column (Altex, C-18, Ultra-phase ion-pairing column; 5 μ m particle size, 25 $\times$ 0.46 cm) was eluted (1 ml min⁻¹) at 22 $^{\circ}$ C with an isocratic mobile phase composed of 29.7% (v/v) acetonitrile prepared in an aqueous buffer containing 0.10 M phosphoric acid, 0.02 M triethylamine and 0.05 M NaClO₄, adjusted to pH 3.0 with NaOH; the instrument was a Perkin-Elmer Series 4 liquid chromatograph equipped with an ISS-100 autosampler and an LC 65-T detector.

Natural and semisynthetic insulins differing by single amino acid residues served as models for examining HPLC-based insulin separations. As shown in Fig. 1*a*, insulins from rabbit, man and pig (having serine, threonine and alanine, respectively, at position B30, but being otherwise identical) elute from the reverse-phase column in the predicted order of decreasing hydrophilicity. Substitution of Phe^{B24} or Phe^{B25} by leucine in human insulin (Fig. 1*b*) or porcine insulin (Fig. 1*c*), however, results in less easily interpreted alterations in peptide elution: in both cases, the insulins with a Leu^{B25} substitution show decreased retention, whereas those with a Leu^{B24} substitution show increased retention, relative to the respective natural hormones. Conformational changes (reflected in the altered interaction of more than one side chain with the hydrophobic column) probably account for the delayed elution of the Leu^{B24}-substituted pig and human insulin (see refs 5–10). Figure 1*d* shows a chromatogram of the eight natural and semisynthetic insulins. The co-elution of rabbit insulin (analogous to porcine insulin with an amino acid substitution at position B30) and porcine [Leu^{B25}]insulin emphasizes the need for caution in assigning elution differences to specific amino acid changes, but the resolving power of the system is clear.

Radioimmunoassay of insulins separated by HPLC showed that recoveries are high. Figure 2*a*, profile 1, shows that peak broadening is minimal (despite the need for collecting fractions before analysis) and that relative elution position is unchanged even when examining fmol quantities of human insulin and

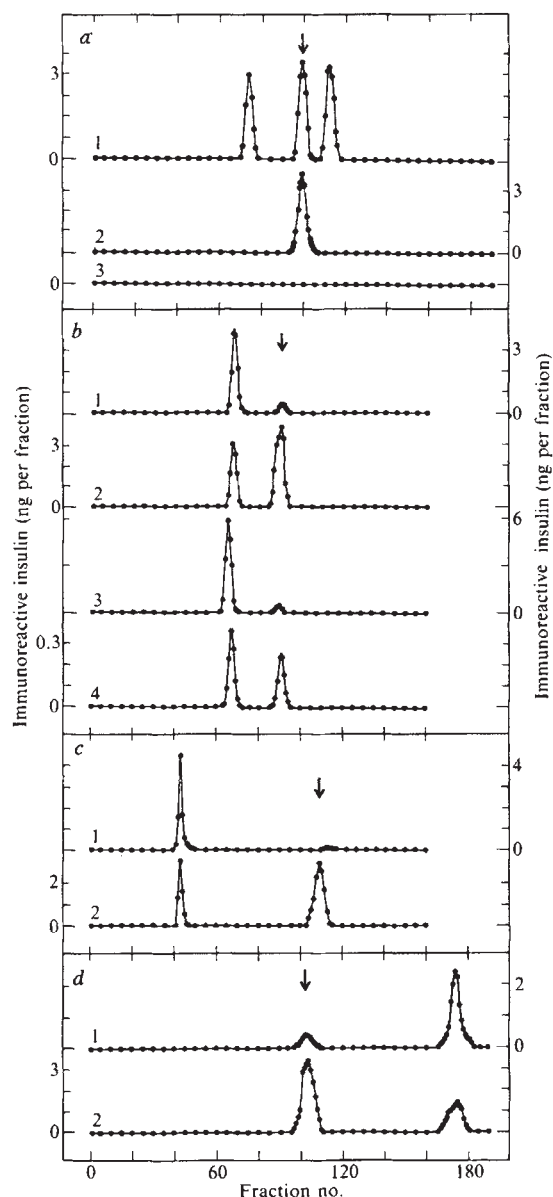


Fig. 2 Separation of serum insulin by reverse-phase HPLC. *a*, Controls: 1, a mixture of human [Leu^{B25}]insulin, natural human insulin and human [Leu^{B24}]insulin (20 ng of each); 2, insulin isolated from the pooled sera of normal subjects; 3, injection of 0.5 mg of bovine serum albumin with no added insulin. *b*, Insulin from patient 1: 1, serum insulin; 2, serum insulin plus added normal human insulin; 3, serum insulin plus semisynthetic human [Leu^{B25}]insulin; 4, pancreatic insulin (3 ng) isolated and described previously¹. *c*, Insulin from patient 2: 1, serum insulin; 2, serum insulin plus normal human insulin. *d*, Insulin from patient 3: 1, serum insulin; 2, serum insulin plus normal human insulin. The position of normal insulin is indicated by the vertical arrow in each panel. Ordinate scales alternate between left and right for clarity of presentation.

Methods: All serum insulin samples were purified from 5–20 ml of serum by immunoaffinity chromatography^{2,14} before HPLC analysis. In each case, gel filtration showed that 90% of these insulin-related peptides had the molecular weight of insulin; proinsulin elutes from the HPLC column only at a much higher concentration of acetonitrile. Serum insulin (20 ng) and standards (20 ng, where indicated) were dissolved in 0.1 ml of 3 M acetic acid containing 0.5 mg of bovine serum albumin for injection and subsequent chromatography as described in the legend to Fig. 1. Fractions (0.4 ml) were collected in tubes containing a drop of borate buffer (0.5 M boric acid plus 10 mg per ml bovine serum albumin brought to pH 9.3 with NaOH) and were dried under vacuum. The residues were dissolved in radioimmunoassay buffer¹⁵ and aliquots were removed for insulin radioimmunoassay¹⁵. Although all fractions were assayed, data are shown only for every fifth fraction when insulin content was below the limit of detection.

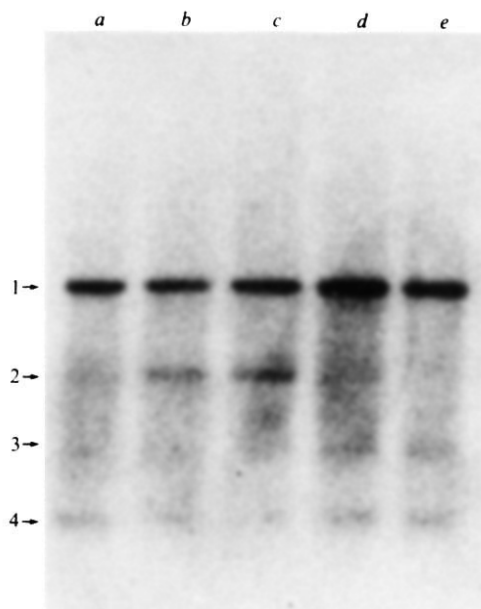


Fig. 3 Restriction endonuclease *Mbo*II mapping of chromosomal insulin genes. *a*, *e*, DNA from two normal subjects. *b*, *c*, *d*, DNA from patients 1, 2 and 3, respectively. Bands 1, 2, 3 and 4 represent 1,600, 920, 580 and 340 bp, respectively. Bands 1, 3 and 4 are characteristic of normal insulin genes, whereas band 2 is found only in insulin genes modified by mutation at the usual *Mbo*II cleavage site (TCTTC) (ref. 3 and unpublished results.) Leukocyte DNA from patient 1 was the gift of Dr S. C. M. Kwok; Leukocyte DNA from other patients and the normal subjects was isolated as described previously³. Each DNA sample (30 µg) was digested with *Mbo*II (90 units, New England Biolabs) for 14 h at 37 °C; a second addition of the enzyme (90 units) and reincubation at 37 °C for 8 h ensured complete digestion. The DNA fragments were separated on a 2% agarose gel and were then transferred to a nitrocellulose filter by the method of Southern¹⁶. The filter was then incubated with cloned human insulin cDNA³ (the gift of Dr S. J. Chan) which had been ³²P-labelled by nick-translation¹⁷, washed extensively, dried and exposed to X-ray film at -70 °C for 7 days with an intensifying screen. ΦX174 DNA digested with *Hae*II was used as a molecular weight standard. (The faintness of bands 3 and 4 in all cases is due to incomplete binding of the corresponding small DNA fragments to nitrocellulose filters.) See text for further details.

human insulin analogues. Figure 2*a*, profile 2, shows that the methods used (including immunoaffinity purification of serum insulin, HPLC separation and radioimmunoassay) can identify the hormone from the serum of normal subjects; profile 3 shows an albumin control.

Our analysis of insulin from patient 1 (the patient previously identified as R.C. and previously shown to secrete an abnormal pancreatic insulin having a leucine for phenylalanine substitution at position B24 or B25)¹ is illustrated in Fig. 2*b*. Profile 1 shows that serum insulin from this patient consists predominantly of an abnormal form which elutes before normal human insulin; small amounts of normal insulin (about 5% of the total) were also present. The abnormal form was clearly separated from normal insulin when the two were mixed (profile 2) and appeared at the position taken by the semisynthetic human [Leu^{B25}]insulin standard (see Fig. 2*a*). Chromatography of the patient insulin with the semisynthetic standard resulted in a single peak identical to [Leu^{B25}]insulin (profile 3). Figure 2*b*, profile 4, shows that the pancreatic insulin from this patient also consists of [Leu^{B25}]insulin and normal human insulin; quantitative analysis showed, as expected¹, that the variant form represented 57% and the normal form 43% of the patient's stored pancreatic insulin.

Studies on the serum insulins from patients 2 and 3 are shown

in Fig. 2*c* and *d*, respectively. The sera from patients 2 and 3 also contained small amounts of normal insulin (about 4% and 9% of the total in the two patients, respectively) and much larger amounts of abnormal insulin (profiles 1 of Fig. 2*c*, *d*). A comparison of profiles 1 of Fig. 2*b*, *c* and *d* shows, however, that the elution positions of the variant insulins from the three patients are very different and that each of the variant insulins must result from a unique amino acid substitution. Thus, the variant insulin from patient 2 elutes well before the least hydrophobic standard (porcine [Ala^{B24}]insulin), the variant from patient 1 co-elutes with human [Leu^{B25}]insulin, and the variant from patient 3 elutes well after the most hydrophobic standard (porcine [Leu^{B24}]insulin).

The availability of leukocyte DNA from each of the patients identified above permitted a further test of abnormality at the level of the insulin gene. Restriction mapping of the insulin gene from patient 1 has already documented loss of a restriction endonuclease *Mbo*II recognition site in one of the patient's insulin gene alleles³. Figure 3 compares the products of *Mbo*II endonuclease digestion of chromosomal DNA isolated from two normal subjects and from patients 1, 2 and 3. As shown in the first and last lanes of the Southern blot, DNA from the normal subjects was cleaved into three major bands corresponding to 1,600, 580 and 340 base pairs (bp) (bands 1, 3 and 4, respectively). In contrast, *Mbo*II digestion of the DNA from patients 1 and 2 resulted in a gene fragment corresponding to 920 bp (band 2) in addition to the three fragments already identified in normal individuals. DNA from patient 3 gave only the three fragments characteristic of normals. As previously described, the 920-bp fragment corresponds to the sum of the 580- and 340-bp genomic fragments and arises from a nucleotide base change in the *Mbo*II-specific sequence TCTTC³. We ascribe band 1 to both insulin alleles, band 2 to the abnormal allele and bands 3 and 4 to the normal insulin allele in each of patients 1 and 2. It thus appears that one insulin gene allele from patients 1 and 2 has undergone mutation at the *Mbo*II-sensitive site corresponding to insulin residues Phe^{B24} and Phe^{B25} (TTC-TTC); the site of mutation in the insulin gene from patient 3 remains unknown.

Our combined use of HPLC and radioimmunoassay has resulted in the identification of three different insulin variants in the sera of three unrelated, hyperglycaemic, hyperinsulinaemic patients who respond normally to exogenously administered insulin. Insulin from patient 1 has already been examined by several approaches^{1,2,11}. Although biological studies involving both the natural patient insulin and semisynthetic Leu^{B24} and Leu^{B25} analogues had suggested that the patient's variant insulin was substituted at position B24, some of these findings could not be reproduced^{12,13} and it is clear that the original assignment is in error. Experiments reported here agree with the alternative proposal of Kobayashi *et al.*¹³, and identify the patient's variant insulin as human [Leu^{B25}]insulin. Our confirmation that the pancreatic insulin from patient 1 consists of [Leu^{B25}]insulin and normal insulin in the molar ratio 60:40 (Fig. 2*b*, profile 4) and our finding that serum insulin from the patient consists of [Leu^{B25}]insulin (an analogue with only 1–2% of normal biological activity^{5–9,12,13} and normal insulin in the molar ratio 95:5, suggests a possible explanation for the patient's hyperglycaemia: as the patient had total, fasting insulin levels of 80–120 µunits per ml of serum², his biologically effective serum insulin concentration was only about 5 µunits per ml, a value below the range seen in many normal individuals. It is probable that a slow clearance of the abnormal insulin from the circulation in this patient results in the high total serum insulin levels observed, and this is also the case for patients 2 and 3.

We do not know the structures of the insulin variants from patients 2 and 3 but, as (1) we have localized the insulin gene mutation in patient 2 to the *Mbo*II-sensitive sequence corresponding to positions Phe^{B24} and Phe^{B25}, (2) single base changes in either codon could result in substitution of phenylalanine by leucine, isoleucine, valine, tyrosine, cysteine or serine, and (3)

the patient's insulin is considerably more hydrophilic than normal insulin, it is most likely that a serine (or cysteine) substitution has occurred at position B24 or B25 in the variant hormone. Further investigation will be required to identify the mutation occurring in the abnormal insulin gene of patient 3, notwithstanding the probable substitution of a hydrophilic residue for one considerably more hydrophobic in the patient's abnormal insulin. Our chemical and genetic methods thus complement each other in identifying patients with mutations within coding regions of the insulin gene. All three patients are heterozygous for an abnormal insulin gene and in all three cases both normal and variant insulin genes are expressed. In addition, our findings suggest that variant human insulins associated with diabetes can arise from mutations at different loci within the insulin gene and can contain a variety of amino acid substitutions. Although

we cannot predict the frequency of mutation within the human insulin gene nor the frequency with which such mutation would be clinically recognizable (many mutations could be phenotypically neutral), it appears that variant products of the human insulin gene may be more common than previously suspected. Using a nomenclature analogous to that developed for abnormal haemoglobins, we propose that the insulins from patients 1, 2 and 3 be named insulins Chicago, Los Angeles and Wakayama, respectively, and that the insulin from patient 1 be designated human insulin B25 (Phe → Leu).

We thank Mrs Anna Mathews for help in preparing the manuscript and the National Diabetes Research Interchange for providing human donor pancreas. The work was supported by NIH grants AM18347, AM13941, AM13914 and AM20595.

Received 31 December 1982; accepted 7 February 1983.

1. Tager, H. *et al.* *Nature* **281**, 122–125 (1979).
2. Given, B. D. *et al.* *New Eng. J. Med.* **302**, 129–135 (1980).
3. Kwok, S. C. M. *et al.* *Biochem. biophys. Res. Commun.* **98**, 844–849 (1981).
4. Haneda, M. *et al.* *Diabetes* **31**, (Suppl. 2), 4A (1982).
5. Tager, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 3181–3185 (1980).
6. Gattner, H. G., Danho, W., Behn, C. & Zahn, H. *Hoppe-Seyler's Z. physiol. Chem.* **361**, 1135–1138 (1980).
7. Inouye, K. *et al.* *Experientia* **37**, 811–813 (1981).
8. Wolmer, A. *et al.* *Hoppe-Seyler's Z. physiol. Chem.* **362**, 581–592 (1981).
9. Inouye, K., Watanabe, K., Tochino, Y., Kobayashi, M. & Shigeta, Y. *Biopolymers* **20**, 1845–1858 (1981).

10. Assoian, R. K., Thomas, N. E., Kaiser, E. T. & Tager, H. S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5147–5151 (1982).
11. Olefsky, J. M., Saekow, M., Tager, H. & Rubenstein, A. H. *J. biol. Chem.* **255**, 6098–6105 (1980).
12. Keefer, L. M. *et al.* *Biochem. biophys. Res. Commun.* **100**, 1229–1236 (1981).
13. Kobayashi, M. *et al.* *Biochem. J.* **206**, 597–603 (1982).
14. Robbins, D. C. *et al.* *Nature* **291**, 679–681 (1981).
15. Starr, J. I., Horwitz, D. L., Rubenstein, A. H. & Mako, M. E. *Methods of Hormone Radioimmunoassay* (eds Jaffe, B. M. & Behrman, H. L.) 613–629 (Academic, New York, 1979).
16. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
17. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).

Formation of proinsulin by immobilized *Bacillus subtilis*

Klaus Mosbach*, Staffan Birnbaum*, Kim Hardy†, Julian Davies† & Leif Bülow*

* Pure and Applied Biochemistry, Chemical Center, University of Lund, PO Box 740, S-220 07 Lund, Sweden

† Biogen S.A., PO Box CH-1211, Geneva 24, Switzerland

There has been an increasing interest in the use of immobilized cells for the production of pharmaceuticals as well as for products such as high fructose syrup or ethanol¹. Some of these compounds are now produced on an industrial scale² whereby the cells are used in a resting or growing state or in a nonviable form as natural carriers of the enzyme(s) involved in the synthesis. The advantages of immobilized cell technology should also apply to microorganisms modified by recombinant DNA techniques to produce a variety of eukaryotic proteins such as hormones. We describe here the properties of immobilized *Bacillus subtilis* cells carrying plasmids encoding rat proinsulin. Cell proliferation normally coupled to DNA replication is undesirable in immobilized cell systems as 'clogging' of the system occurs due to cells growing outside the beads. Therefore, different ways were investigated to inhibit cell division while allowing continued protein synthesis. We found that the addition of certain antibiotics in the growth medium, such as novobiocin which inhibits DNA replication³, fulfills these requirements, allowing proinsulin synthesis and excretion to take place over a period of several days.

The strain used for these experiments was *B. subtilis* BS273 which comprises the host strain SL438 (provided by P. Pigot) carrying plasmid pPCB6. Plasmid pPCB6 was isolated by inserting a fragment specifying rat preproinsulin⁴ into a 4.3-kilobase sequence encoding the penicillinase of *B. licheniformis* which was cloned on plasmid ppen4300⁵ (provided by H. Schaller). A 350-base pair DNA fragment encoding rat preproinsulin was obtained by digesting plasmid p218.CB6 with the restriction endonuclease *Pst*I. This fragment was inserted at the *Pst*I site which occurs within the DNA sequence encoding the NH₂-terminal signal peptide of *B. licheniformis* penicillinase⁵. Ligation of the DNA fragments and selection of the recombinant plasmids by transformation were done using methods previously described⁶. Plasmid pPCB6 thus encodes the first 15 amino

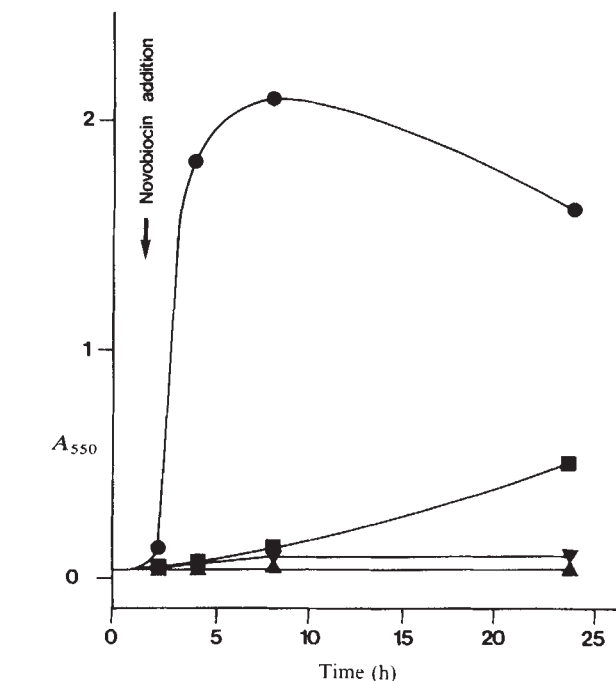


Fig. 1 The effect of novobiocin on growth of free *B. subtilis* BS273 cells. The cells were grown at 37 °C with vigorous shaking (150 r.p.m.) in L-broth supplemented with 12.5 µg ml⁻¹ tetracycline. Novobiocin concentration used were 0 µg ml⁻¹ (●), 2 µg ml⁻¹ (■), 5 µg ml⁻¹ (▼), and 8 µg ml⁻¹ (▲).

acids of the penicillinase signal followed by the signal sequence of rat preproinsulin.

A number of support materials including alginate, polyacrylamide and agarose were tested for their ability to entrap the cells as well as to release proinsulin. Agarose 2% (w/v) was most effective as it allowed rapid release of entrapped ¹²⁵I-labelled insulin; in addition it is non-toxic. Agarose beads (2 mm diameter) released 80% of the entrapped ¹²⁵I-insulin after 2 hours' incubation in 0.1 M Tris-HCl (pH 7.5) whereas polyacrylamide (10% w/v) or calcium alginate (2% w/v) beads released only 15% and 20% of the entrapped insulin. The concentration of novobiocin required to arrest growth of *Bacillus subtilis* in suspension cultures was tested and it was found (Fig. 1) that at a concentration of 5 µg ml⁻¹, growth was

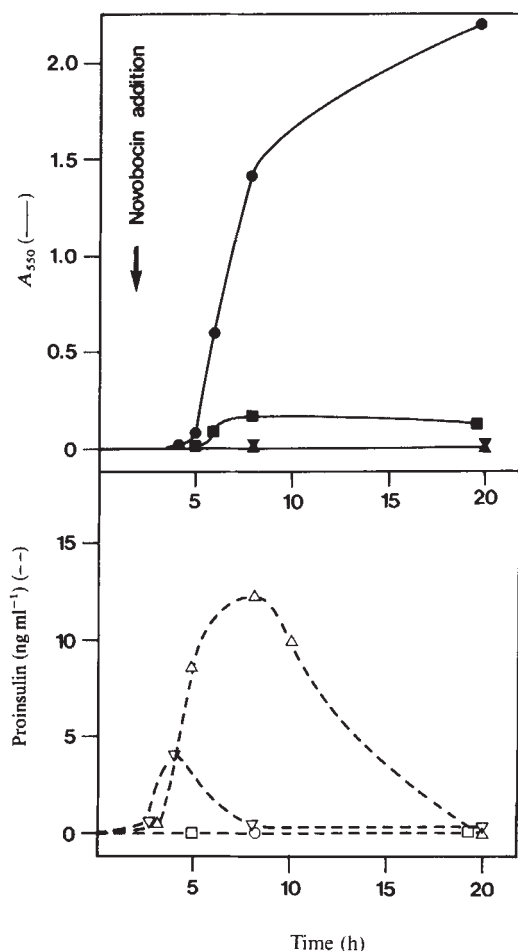


Fig. 2 *B. subtilis* cells were grown in 25 ml medium as described in Fig. 1 and collected during exponential growth at A_{550} 1.0. The cells were centrifuged (5,000g) and washed twice with phosphate-buffered saline (PBS). Twenty-five ml of the cell suspension was mixed with an equal volume of melted 4% (w/v) agarose (type VII, Sigma), kept at 42 °C, and poured into a 1-litre beaker containing 500 ml soya oil (prewarmed to 37 °C) under vigorous stirring. When the formed beads were 100–300 μm in diameter the suspension was cooled to below the setting temperature of agarose. After the beads had solidified, 300 ml PBS were added, stirring ceased, and the immobilized preparation was allowed to settle overnight at room temperature. Subsequently, the oil layer was decanted and discarded (the final droplets were removed with a pipette). Finally, the beads were filtered on a nylon net (250 mesh) and washed with PBS. The immobilized *B. subtilis* preparation (5 g, wet weight) was added to 20 ml of the above medium in 100 ml Erlenmeyer flasks. The flasks were incubated on a rotary shaker (150 r.p.m.) at 37 °C. The A_{550} and the proinsulin content of the medium were determined by removing 1 ml samples. After incubation for 2 h novobiocin was added to the medium. The rat proinsulin produced was quantified by liquid radioimmunoassay⁴. Proinsulin formation (-----) was determined at different novobiocin concentrations: 0 $\mu\text{g ml}^{-1}$ (○), 2 $\mu\text{g ml}^{-1}$ (□), 5 $\mu\text{g ml}^{-1}$ (▽), and 8 $\mu\text{g ml}^{-1}$ (△). The corresponding absorbances (—) were determined at 550 nm.

Table 1 Cell growth inside agarose beads with and without addition of novobiocin

Hours	Cell number per g immobilized beads (wet weight)	
	Without novobiocin	With novobiocin
0	2×10^8	—
2	2.2×10^8	—
5	4.2×10^8	4.3×10^8
20	8.9×10^8	4.2×10^8

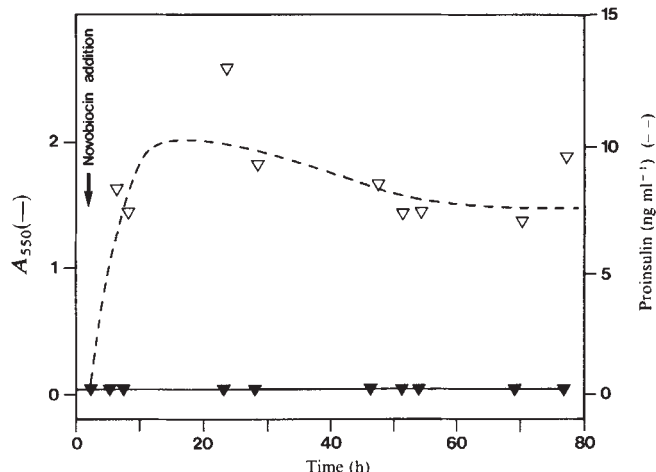


Fig. 3 Continuous formation of proinsulin by immobilized *B. subtilis*. Five grams of the immobilized preparation were added to 20 ml of the medium in 100 ml Erlenmeyer flasks and shaken (150 r.p.m.) at 37 °C. Novobiocin (5 $\mu\text{g ml}^{-1}$) was added to the medium after incubation for 2 h. Fresh medium supplemented with novobiocin was continuously added to the small continuous stirred tank reactor at a flow rate of 4 ml h^{-1} . The A_{550} (▼) and the amount of proinsulin (▽) in the effluent were determined by taking aliquots of 1 ml.

inhibited but a similar amount of proinsulin (7–10 ng ml^{-1}) was found in the medium as in the untreated control. The *B. subtilis* cells were then included in agarose and formed into minute beads (100–300 μm diameter) using a method in which inert oil acts as suspension medium^{7,8}. The viability of the entrapped cells based on cell growth and oxygen consumption was tested. As seen from Table 1 the number of cells inside the agarose beads increased over a period of 20 h. On addition of 5 $\mu\text{g ml}^{-1}$ novobiocin after 2 h of growth, cell division was completely arrested. These figures were obtained by correlating the total protein content found in the immobilized preparations to that found with free cells⁹.

Oxygen consumption of the immobilized preparations was also determined. On average about 0.4–0.8 μg per min per g bead preparation (wet weight) was consumed as determined with an oxygen electrode¹⁰. These measurements were carried out by washing the immobilized preparation on a glass filter and adding 0.1 g of the preparation to 2.9 ml air-saturated medium and oxygen consumption measured. No significant difference was found for the different preparations after growth for 5 and 20 h respectively.

Immobilized *B. subtilis* cells were incubated in L-broth medium supplemented after 2 h of incubation with various amounts of novobiocin. As seen from Fig. 2, at concentrations of 5 or 8 $\mu\text{g ml}^{-1}$ no cell growth was found but proinsulin production continued. The fact that the proinsulin levels decreased is probably due to protease action. In the control (no novobiocin added) considerable cell growth was found outside the beads. (Again proteases presumably account for the fact that no proinsulin was detected.)

We then carried out a study on the continuous production of proinsulin using immobilized *B. subtilis* cells in a small model for a continuous stirred tank reactor. Although novobiocin prevented cell growth outside the beads, continuous formation of proinsulin over a period of 80 h took place (Fig. 3). The amount of proinsulin formed was constant in the range 7–10 ng per ml of medium (9,000–13,000 molecules per cell) as determined by radioimmunoassay⁴.

This effect of novobiocin, acts by inhibiting supercoiling of DNA catalysed by DNA gyrase and thus DNA replication³ whilst not affecting protein synthesis, indicates the potential usefulness of such an approach to the production of excreted protein by immobilized cells obtained by recombinant DNA techniques. In a preliminary study using the same approach with *B. subtilis* producing α_2 -interferon¹¹ it could be shown

that in the presence of novobiocin, cell growth outside the beads was prevented while continuous interferon excretion into the medium took place.

Other antibiotics tested, including nalidixic acid and 6-(p-hydroxyphenylazo)-uracil had similar effects. Thus, as tested with a genetically engineered strain of a periplasmic leaky mutant of *E. coli* (EC623; S. Stahl and K.H., unpublished results), addition of 0.1 mM nalidixic acid leading to the inhibition of cell growth, still maintained a production level of proinsulin of 100 ng ml⁻¹ found in the medium. Similar results were obtained with free cells. Thus the technique of arresting cell growth by inhibition of cell division with antibiotics while protein biosynthesis/excretion continues should also be useful for conventional fermentation processes using free cells as lesser amounts of cells are required for the same amount of hormones formed.

In the case of immobilized systems using genetically engineered organisms, arresting of cell growth, however, appears to be a prerequisite for hormone production. Provided that protein synthesis can be maintained over long periods of time while controlling cell growth (with novobiocin or nalidixic acid, for example), immobilized systems offer a valuable alternative to conventional fermentation techniques using cells in suspension. Advantages of this technique include: high cell densities in a dilute system; the cells can be protected from damage (possibly also leading to delayed lysis); and no removal of the product from the cells is required. As a further step the on-line coupling of specific affinity columns to an immobilized system can be envisaged to allow rapid isolation of the product.

The valuable contribution by Ms E. Lohmeier is acknowledged. This investigation has in part been supported by the National Swedish Board for Technical Development (803595).

Received 13 December 1982; accepted 8 February 1983.

1. Mosbach, K. (ed.) *Methods in Enzymology* 44 (Academic, New York, 1976).
2. Chibata, I. in *Food Process Engineering* Vol. 2 (eds Linko, P. & Larinkari, J.) 1-26 (Applied Science, London, 1980).
3. Gellert, M., O'Dea, M., Itoh, T. & Tomizawa, J.-I. *Proc. natn. Acad. Sci. U.S.A.* 73, 4474-4478 (1976).
4. Talmadge, K., Stahl, S. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* 77, 3369-3373 (1980).
5. Neugebauer, K., Sprengel, R. & Schaller, H. *Nucleic Acids Res.* 9, 2577-2588 (1981).
6. Hardy, K., Stahl, S. & Kupper, H. *Nature* 293, 481-483 (1981).
7. Wikström, P., Szwajcer, E., Brodelius, P., Nilsson, K. & Mosbach, K. *Biotechnol. Lett.* 4, 153-158 (1982).
8. Nilsson, K. & Mosbach, K. Swedish Patent Application No. 8210041-0.
9. Freeman, A., Blank, T. & Aharonowitz, Y. *Eur. J. appl. Microbiol. Biotechnol.* 14, 13-15 (1982).
10. Brodelius, P. & Nilsson, K. *FEBS Lett.* 122, 312-316 (1980).
11. Palva, I. *et al. Gene* (in the press).

Inhibition of adenovirus early region IV transcription *in vitro* by a purified viral DNA binding protein

Hiroshi Handa*, Robert E. Kingston† & Phillip A. Sharp†

*The Institute of Medical Sciences, The University of Tokyo, 4-6-1 Shikanedai, Minato-ku, Tokyo, Japan

†Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Adenoviruses depend on cellular mechanisms for the decoding of their genetic information, and so provide a useful and simple model system for the investigation of mammalian gene expression. The five regions transcribed early in adenovirus infection are termed EIa, EIb, EII, EIII and EIV^{1,2}. We report here that the primary product of the EII region, a 72,000 molecular weight DNA-binding protein³ (DBP), specifically represses transcription from the EIV promoter in an *in vitro* transcription system. Single-stranded DNA binds to the DBP with high affinity, and as a result inhibits its repressor activity. Our data extend previous genetic evidence that the DBP represses EIV transcription *in vivo*⁴⁻⁶, and suggest that it acts directly by suppressing transcription from the EIV promoter.

Adenovirus strain H5 *ts125*, which encodes a temperature-sensitive DNA binding protein, has been used to study the function of this protein *in vivo*³⁻⁶. Carter and Blanton have shown that levels of cytoplasmic RNAs complementary to all five early transcription units are increased in H5 *ts125*-infected cells at the non-permissive temperature relative to the permissive temperature⁵. More recently, Nevins and Winkler showed by pulse-labelling during H5 *ts125* infections that EIV transcription was repressed at 6 h post-infection at the permissive temperature, but not at the non-permissive temperature⁶. This latter study, which used high multiplicities of infection, suggested that the DBP represses transcription, only from the EIV region. Earlier studies showed that DBP binds to single-stranded DNA in a cooperative manner, can also bind to the termini of double-stranded DNA, and is essential for viral DNA replication^{3,7-9}.

Whole cell extracts prepared from uninfected HeLa cells are capable of initiating transcription specifically *in vitro* at several adenovirus promoters^{10,11}. We have compared the effects of purified DBP on transcription initiating at the EIa, EIb, EIV and major late (MLP) promoters. Recombinant plasmids containing the adenovirus promoters were cleaved with the appropriate restriction endonuclease so that transcription would produce 'run-off' transcripts of defined lengths. Run-off transcripts span the length between the point of transcription initiation and the end of the restriction fragment.

Preincubation of viral DNA fragments with active DBP before addition of the extract preferentially represses transcription from the EIV promoter. Mixtures of DNA fragments were preincubated either with control buffer or with purified DBP, and then diluted with a properly constituted whole cell extract reaction mix. After a 60 min incubation in the presence of ³²P-labelled UTP, the RNA products were extracted and resolved by electrophoresis in denaturing glyoxal gels. Autoradiographs of gels such as that shown in Fig. 1a were traced using a densitometer, and the amount of radioactivity migrating in various bands was compared. In general, preincubation with high concentrations of DBP inhibited transcription from all viral promoters; however, at lower concentrations of DBP, specific suppression of transcription of the EIV promoter site was observed (Fig. 1a).

To quantitate the level of specific inhibition of transcription from the EIV promoter, the amount of radioactivity in each band was normalized to that transcribed from the MLP fragment included in the same reaction mixture. This comparison is plotted in Fig. 1b and c as a function of the molar ratio of DBP:template DNA. The quantitation confirms the visual impression that transcription of the EIV promoter is specifically suppressed relative to the MLP, EIa or EIb promoters.

As total transcription was inhibited by addition of DBP, it was necessary to control for a possible effect of the length of run-off transcripts on specific suppression. Run-off transcripts of lengths 680, 1,220 and 2,010 nucleotides from EIV were specifically suppressed relative to transcripts of greater length from the MLP. Inhibition by 1 µg ml⁻¹ α-amanitin¹² showed that the transcription was the product of class II RNA polymerase. Heating the DBP for 5 min at 60 °C before addition to the reaction mixture eliminated the selective inhibition of the EIV promoter.

The DBP is so named because it binds in a cooperative fashion to single-stranded DNA. This specific binding provided another means of testing whether the selective inhibition was due to the activity of the DBP. Sufficient DBP was preincubated with a mixture of viral fragments to produce a 60% inhibition specific to the EIV promoter. The effect of including increasing amounts of single-stranded DNA on this inhibition was determined in parallel reactions. Transcription from the mixture of promoter fragments was measured as described previously. As shown in Fig. 2, inhibition of transcription of EIV was significantly reduced by the addition of single-stranded DNA. Addition of single-stranded DNA did not alter the relative level of transcription from either the EIa or EIb promoters. These

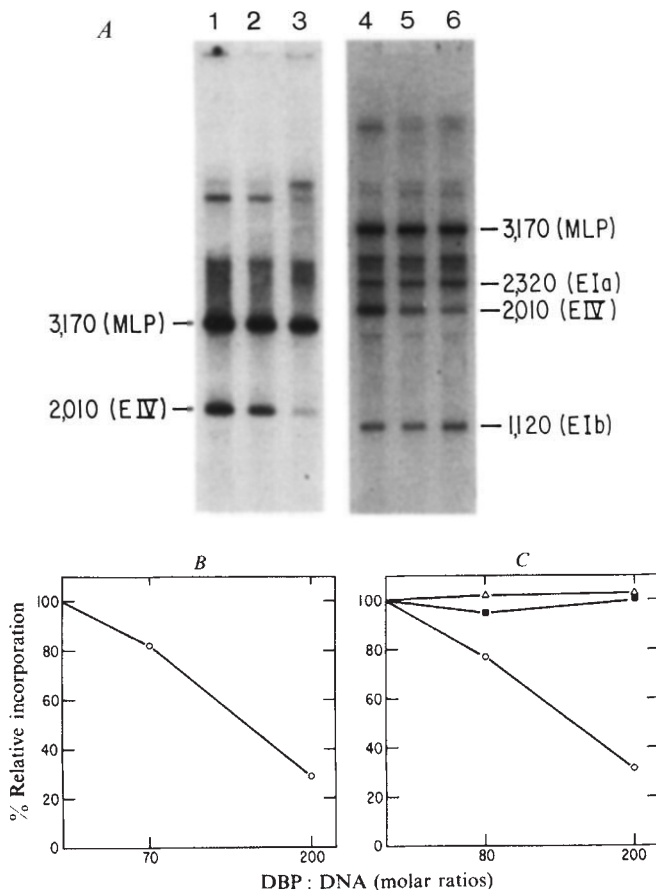


Fig. 1 The effect of DBP on transcription from adenovirus promoters *in vitro*. **A** Shows the labelled transcripts made in the system described below, with the following template DNAs: lanes 1–3, *KpnI*-cleaved pRIB5 (map units 84–100 of Ad2 which contain the EIV promoter), inserted into the pBR322 (*EcoRI* site) and *AvaI*-cleaved pFLBH (map units 14–17 of Ad2, which contain the MLP, inserted between the *HindIII* and *BamHI* sites of pBR322); lanes 4–6, the same as lanes 1–4, but the reactions also contained *HindIII*-cleaved p*HindIII*G (map units 0–7.9 of Ad5, which contain the E1a and E1b promoters, inserted into the *HindIII* site of pBR322). The molar ratios of DBP: template DNA were as follows: lanes 1 and 4, no DBP; lane 2, 70:1; lanes 3 and 6, 200:1; lane 5, 80:1. The RNA transcripts are marked according to their length in nucleotides, and are identified by promoter as determined previously¹¹. **B** Shows a graph obtained by quantitation of lanes 1–3 of the autoradiograph shown in **A**; and **C** shows a similar graph from quantitation of lanes 4–6. The intensity of the bands in **A** was quantitated by tracing the autoradiographs with a Zeineh laser densitometer. The amounts of the E1a, E1b and EIV bands were calculated relative to the MLP band, this value at no added DBP being assigned as 100%. The change in this relative value is plotted as a function of increasing DBP. The level of MLP transcript in each lane varied as follows: lane 2, 81% of lane 1; lane 3, 68% of lane 1; lanes 5 and 6, 81% of lane 4. ○, EIV transcript; △, E1b transcript; ■, E1a transcript.

Methods: The plasmids were provided by Kathleen Berkner and Frank Laski, and are described elsewhere¹¹. Adenovirus DBP was purified from Ad5-infected HeLa cells by the method of Ikeda *et al.*¹⁵. DBP and the template DNA (10–15 $\mu\text{g ml}^{-1}$) were combined at the indicated molar ratios, and incubated for 15 min at room temperature. Whole cell transcription extract¹⁰ was mixed with poly[d(IC)-d(IC)] and cofactors on ice, and then added to the template DNA and DBP. Transcription reactions were incubated for 60 min at 30 °C. RNA was extracted from the reaction mixture and analysed by glyoxalation and agarose gel electrophoresis as described^{10,11}. A standard complete reaction contained 10 mM HEPES, pH 7.9, 40 mM KCl, 5 mM MgCl₂, 0.05 mM EDTA, 0.8 mM dithiothreitol, 7% glycerol, 10 mM creatine phosphate, 200 μM ATP, 100 μM CTP and GTP, 5 μM UTP, [α -³²P]UTP (400 $\mu\text{Ci ml}^{-1}$), 18 $\mu\text{g ml}^{-1}$ poly[d(IC)-d(IC)], 4–7 $\mu\text{g ml}^{-1}$ template DNA and HeLa cell extract.

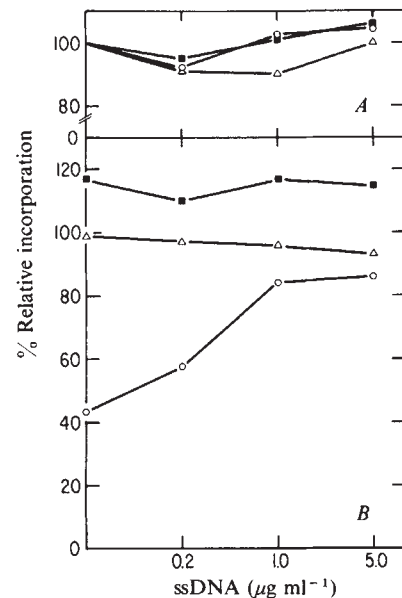


Fig. 2 Release of inhibition of EIV transcription by the addition of single-stranded DNA. Reactions were carried out and the product RNA was analysed exactly as described for lanes 4–6 of Fig. 1A. Sonicated calf thymus DNA was denatured by heat and included in the DNA template/DBP preincubation reaction at the indicated concentrations. Autoradiographs were traced with a densitometer and the intensities of the E1a, E1b and EIV bands were determined relative to the MLP band (this value, in the absence of added single-stranded DNA (ssDNA) and DBP was assigned as 100%). The relative intensities of the E1a, E1b and EIV products are plotted as a function of ssDNA concentration. **A**, No DBP added; **B**, DBP included at a molar ratio to template DNA of 200:1. ○, EIV transcript; △, E1b transcript; ■, E1a transcript.

results further support the conclusion that transcription from EIV is specifically repressed by DBP.

Infection of HeLa cells at non-permissive temperatures by Ad5 *ts125* results in the over-production of early mRNAs from EIV⁵. Following the rate of transcription by short pulses of ³H-uridine suggested that this over-production is due to the lack of inhibition of transcription from EIV by some activity of the DBP⁶. Our experiments suggest that the DBP acts directly to suppress transcription from EIV. This inhibition is specific for EIV at relatively low ratios of DBP:DNA template. Whether the suppression of transcription from other viral promoters at higher levels of DBP reflects an *in vivo* process is not clear.

Only one case of regulation of RNA polymerase II transcription is understood at the molecular level, autoregulation of SV40 early transcription by its A gene product, the T antigen^{13,14}. In this case T antigen binding is thought to block sterically the binding of RNA polymerase II to the promoter. The DBP may limit initiation of transcription at the EIV promoter by a similar process, or by interacting directly with RNA polymerase. Alternatively, the DBP may act by terminating transcription at a post-initiation stage. By utilizing the *in vitro* assay reported here, it should now be possible to distinguish between these possibilities.

We thank Ulla Hansen, Andrew Fire, Mark Samuels and Kathleen Berkner for their help. H. H. was supported by the USA–Japan Cooperative Cancer Research Program from the Japan Society for Promotion of Science, and R. E. K. received support from the Jane Coffin Childs Memorial Fund. This work was supported by grants PO1-CA26717 from the NIH, grant PCM78-23230 (currently PCM82-00309) from the NSF to P.A.S., and grant PO1-CA14051 from the NIH Center for Cancer Biology at MIT.

Received 17 December 1982; accepted 1 February 1983.

1. Sharp, P. A., Gallimore, P. H. & Flint, S. J. *Cold Spring Harb. Symp. quant. Biol.* **39**, 457-474 (1975).
2. Ziff, E. B. *Nature* **287**, 491-499 (1980).
3. Ensinger, M. & Ginsberg, H. S. *J. Virol.* **10**, 328-338 (1972).
4. van der Vliet, P. C., Levine, A. J., Ensinger, M. & Ginsberg, H. S. *J. Virol.* **15**, 348-354 (1975).
5. Blanton, R. & Carter, T. J. *J. Virol.* **29**, 458-465 (1979).
6. Nevins, J. R. & Winkler, J. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1893-1897 (1980).
7. van der Vliet, P. C. & Levine, A. J. *Nature* **246**, 170-174 (1973).
8. Fowlkes, D. M., Lord, S. T., Linne, T., Pettersson, U. & Philipson, L. *J. molec. Biol.* **132**, 163-180 (1979).
9. Horwitz, M. S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4291-4295 (1978).
10. Manley, J. A., Fire, A., Cano, A., Sharp, P. A. & Gelfand, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3855-3859 (1980).
11. Fire, A., Baker, C. C., Manley, J. L., Ziff, E. B. & Sharp, P. A. *J. Virol.* **40**, 703-719 (1981).
12. Roeder, R. G. in *RNA Polymerase* (eds Losick, R. & Chamberlin, M.) 285-329 (Cold Spring Harbor Laboratory, New York, 1976).
13. Rio, D., Robbins, A., Myers, R. & Tjian, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5706-5710 (1980).
14. Hansen, U., Tenen, D. G., Livingston, D. M. & Sharp, P. A. *Cell* **27**, 603-612 (1981).
15. Ikeda, J.-E., Enomoto, T. & Horwitz, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 884-888 (1981).

Homology between an endogenous viral LTR and sequences inserted in an activated cellular oncogene

Edward L. Kuff, Anita Feenstra & Kira Lueders

Laboratory of Biochemistry, National Cancer Institute,
National Institutes of Health, Bethesda,
Maryland 20205, USA

Gideon Rechavi, David Givol & Eli Canaani

Department of Chemical Immunology, Weizmann Institute of
Science, Rehovot 76100, Israel

Recently, some of us reported¹ the detection and molecular cloning of a rearranged cellular oncogene, designated *rc-mos*, from a non-virally-induced mouse myeloma, XRPC24. Recombinant λ phage DNA containing the *rc-mos* gene was active in transforming NIH 3T3 cells in a transfection assay, whereas recombinant DNA containing the unrearranged *c-mos* gene was not. In *rc-mos*, coding sequences from the 5' end of *c-mos* were found to have been displaced by a novel cellular element whose nucleotide sequence was reported. We now document the fact that a 349-base pair (bp) segment of the novel DNA immediately adjacent to the retained *c-mos* sequences in *rc-mos* has close homology with the long terminal repeat (LTR) of a known intracisternal A-particle gene. This homology was mentioned in *Nature* recently² after it had been brought to the attention of the editors (N. Hozumi and R. Hawley, personal communication).

Intracisternal A-particles (IAPs) are retrovirus-like entities abundantly expressed in many mouse tumour cells, including myelomas³. Endogenous genetic elements (IAP genes) homologous to the high-molecular weight RNA of IAPs are reiterated to the extent of 1,000 or more copies per haploid genome in *Mus musculus*⁴. Several of these elements have been cloned and analysed⁵⁻⁷; both the full-size 7.2-kilobase pair (kbp) genes and shorter units modified by deletions in various positions contain LTR sequences 350-400 bp long⁷⁻⁹. Recently, it was shown that functionally defective κ light-chain genes in two mutant mouse hybridoma sublines contained novel insertions of IAP genetic elements including the LTR regions^{10,11}. Although complete sequence data on the inserted elements themselves are not yet available, both LTRs of a 7.2-kbp IAP gene (MIA14) previously isolated from a mouse embryo genomic library^{5,7} have been fully sequenced and have been shown¹¹ to share many properties with proviral LTRs of known infectious retroviruses¹².

In Fig. 1, the reported sequences of the novel DNA element in *rc-mos* and the 5' LTR of MIA14 have been aligned with the aid of a computer program¹³. *rc-mos* is shown in the correct sense (5' to 3') with respect to transcription of the *c-mos* sequence. The MIA14 LTR is oriented in the direction opposite to that of normal IAP gene transcription, the antisense strand being shown. Thus, in *rc-mos* the LTR-homologous component and the retained *c-mos* coding region are arranged in a head-to-head configuration with respect to each other. Note the short terminal inverted repeats (underlined) and (in base-complementary form) a possible TATA box at positions 186-191 of the MIA14 LTR sequence and a canonical polyadenylation signal at positions 265-270.

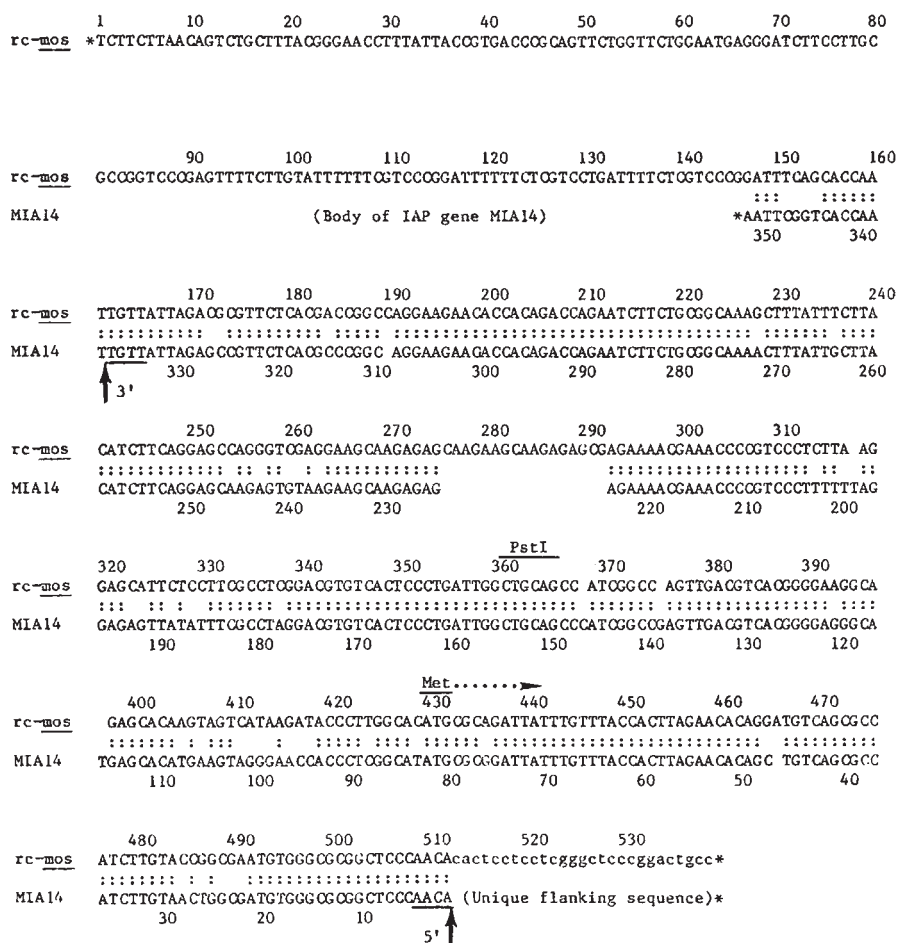
The 336 nucleotides of the MIA14 LTR are homologous in 295 positions (88%) to the *rc-mos* sequence. The junction between the novel DNA and retained *c-mos* at position 512 on the *rc-mos* sequence corresponds precisely to the 5' boundary of the LTR sequence in MIA14. The 28-codon open reading frame previously noted to begin at position 430 in the *rc-mos* sequence¹ is seen to lie within the LTR region but in the opposite sense and different location from the normal IAP transcription. The new in-frame histidine (CAC) codon at the *rc-mos* junction is formed by the last two nucleotides of the LTR inverted repeat sequence and the first retained nucleotide of *c-mos*. The polyadenylation signal found in the MIA14 LTR is located in a corresponding position (230-235) in the *rc-mos* LTR sequence. However, *rc-mos* lacks an apparent TATA box in the position corresponding to the one in MIA14 (or indeed elsewhere).

The LTR in *rc-mos* is 351 nucleotides long (positions 162-512). It contains 37 mismatched bases, 4 gaps and 2 single base pair insertions with respect to the MIA14 LTR. In addition, the *rc-mos* LTR contains a 17-nucleotide segment (positions 276-292) which is absent from the MIA14 sequence. This segment includes a perfect repeat of 12 nucleotides that appear immediately 'upstream' in the *rc-mos* LTR sequence (positions 264-275). Divergence between LTR sequences of individual IAP elements is not unexpected, as these genes have been evolving for several millions years in the lineage of *M. musculus* and are known to be polymorphic in this and other respects⁵⁻⁹.

IAP genes contain an invariant *Pst*I site near the middle of the LTR sequence (see Fig. 1) and, at their 5' ends, a characteristic *Eco*RI site 300 bp further downstream in the body of the gene⁷. In *rc-mos*, position 1 of the known sequence lies in or near an authentic *Eco*RI site at the junction between the cloned mouse genomic fragment and the phage vector arm¹. This site is ~360 nucleotides from the LTR *Pst*I site, that is, in a position appropriate for an IAP gene. MIA14 contains an *Eco*RI site in a similar location^{5,7}, but is unusual in having an additional *Eco*RI site within a few nucleotides of the 5' LTR¹¹. We have not sequenced MIA14 beyond this point. The first 116 nucleotides of the *rc-mos* sequence may represent the 5' region of the body of the inserted IAP genetic element, appearing in complementary form and direction opposite to that of normal IAP transcription. No methionine-initiated open reading frames appear in this sequence.

This is the second time that IAP genetic elements have been shown to appear in novel locations in mouse cellular DNA. In the earlier case, IAP insertions into the introns of κ light-chain genes were associated with defective gene function¹⁰. In the present case, insertion has resulted in the activation of a cellular oncogene. It is not yet known whether the recombination event involved a complete IAP gene or a subgenomic element containing an LTR and some adjoining sequence. A cellular oncogene, *c-myc*, has been shown to be activated in bursal lymphomas by insertion of complete or partial proviral copies after infection by avian leukosis virus (ALV)^{14,15}; and integrated retroviral sequences have also been shown to affect the function of cellular genes involved in mouse coat colour determination¹⁶ and embryonic development¹⁷. Payne *et al.*¹⁵ reported that *c-myc* could be activated by insertion of ALV provirus upstream but in the opposite orientation. In that case, the viral

Fig. 1 The previously reported nucleotide sequences of *rc-mos*¹ and the LTR of the IAP gene MIA14 (ref. 11) have been aligned for optimal homology using the NUCLAN computer program¹³ with the following parameters: k-tuple size = 2, window size = 20, gap penalty = 5. * Indicates the end of sequence data presently available. Vertical arrows indicate the 5' and 3' boundaries of the MIA14 LTR as it is oriented with respect to the whole IAP gene and the customary IAP transcription. The antisense strand of this sequence is shown. The *rc-mos* sequence is presented in the direction of transcription of *c-mos* (sense strand) is shown in lower case letters. Short inverted repeats at the ends of the LTR sequences are underlined. Met at position 430 on the *rc-mos* sequence indicates the beginning of an open reading frame that extends in-phase into the retained *c-mos* region¹.



element was nearby but separated by other sequences from *c-myc*, and fused *c-myc*-viral transcripts were not found in the affected cells. In the present case, the origin of transcription and the means by which it is promoted in the recombinant IAP-*c-mos* sequence are being studied.

The reported examples of IAP sequence insertion have occurred in particle-producing mouse tumour cells. The possibility of new provirus formation and integration in these cells must be considered. Alternatively, the IAP insertions might have resulted from direct transpositions of DNA copies already present in the mouse genome. In either event, the large family of endogenous IAP elements seems to be a significant source of genetic variation in the mouse, and it is possible that IAP genes occasionally have a role in transformation or tumour progression through activation of cellular oncogenes.

We thank Dr Nobumichi Hozumi for bringing to our attention the homology between *rc-mos* and the IAP LTR.

Histone H5 can correctly align randomly arranged nucleosomes in a defined *in vitro* system

A. Stein* & P. Künzler†

* Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA

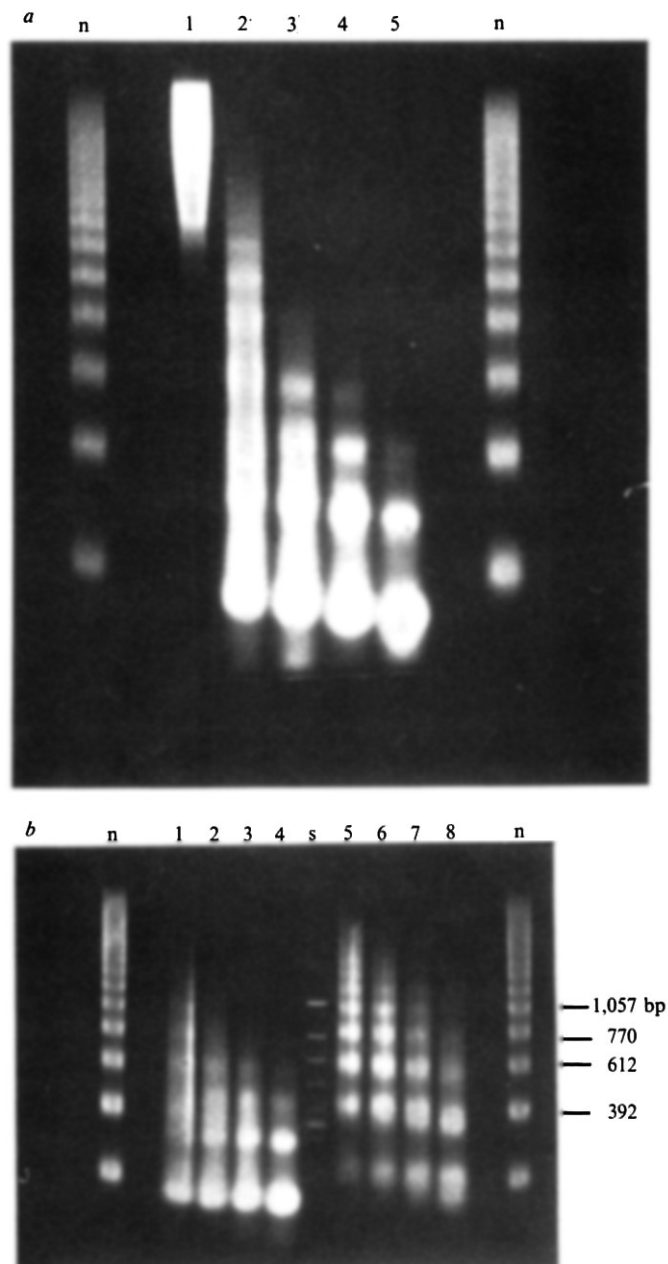
† Institute of General Microbiology, University of Bern, Baltzer-Strasse 4, CH-3012 Bern, Switzerland

In eukaryotic cells, DNA is packed into regularly spaced chromatin subunits called nucleosomes. The average distance between nucleosomes (the repeat length) varies in a tissue- and species-specific manner, with values ranging from about 160 to 240 DNA base pairs (bp)¹. Thus, it has been recognized that the repeat length could be one of the factors underlying selective gene expression. In cells growing in culture, the characteristic repeat length for that type of cell seems to arise from an immature chromatin structure in which nucleosomes are initially irregularly spaced or are arranged in small closely packed clusters²⁻⁵. At present no *in vitro* system has been described which is capable of reconstituting the mature physiological nucleosome spacing from purified chromatin components. Moreover, neither the factors necessary for spacing nor the reaction mechanism are known. We describe here an *in vitro* system that can restore the native subunit spacing in rearranged chromatin samples which have irregularly spaced nucleosomes similar to the situation apparent in newly replicated chromatin.

Received 26 January; accepted 18 February 1983.

- Rechavi, G., Givol, D. & Canaani, E. *Nature* **300**, 607-611 (1982).
- Newmark, P. *Nature* **301**, 196 (1983).
- Dalton, A., Potter, M. & Merwin, R. M. *J. natn. Cancer Inst.* **26**, 1221 (1961).
- Lueders, K. K. & Kuff, E. L. *Cell* **12**, 963 (1977).
- Lueders, K. K. & Kuff, E. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3571 (1980).
- Ono, M., Cole, M. D., White, A. T. & Huang, R. C. *Cell* **21**, 465 (1980).
- Kuff, E. L., Smith, L. A. & Lueders, K. K. *Molec. cell. Biol.* **1**, 216 (1981).
- Cole, M. D., Ono, M. & Huang, K. C. *J. Virol.* **38**, 680 (1981).
- Sheng-Ong, G. L. C. & Cole, M. D. *J. Virol.* **42**, 411 (1982).
- Hawley, R. G., Shulman, M. J., Murialdo, H., Gibson, D. M. & Hozumi, N. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7425 (1982).
- Kuff, E. L. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Varmus, H. E. *Science* **216**, 812 (1982).
- Wilber, W. G. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **80**, 726 (1983).
- Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475-480 (1981).
- Payne, G. S., Bishop, J. M. & Varmus, H. E. *Nature* **295**, 209-214 (1982).
- Jenkins, N. A., Copeland, N. G., Taylor, B. A. & Lee, B. K. *Nature* **293**, 370-374 (1981).
- Jaenish, R. *et al. Cell* **32**, 209 (1982).

Fig. 1 Electrophoresis of DNA fragments produced by micrococcal nuclease digestion of an H1/H5-stripped 'randomized' chromatin sample, untreated (a), or incubated with polyglutamic acid in the absence or presence of H5 (b). Chicken erythrocyte nuclei were prepared as described elsewhere⁹. To prepare chromatin, nuclei containing ~50 mg of DNA were, after two washes, suspended in 10 ml of 0.1 M NaCl, 10 mM Tris-HCl pH 8.0, 1.0 mM EDTA, equilibrated at 37 °C, adjusted to 2 mM CaCl₂, then digested with 50 units of micrococcal nuclease for 4 min. The digestion was stopped by adding 0.5 M EDTA to 10 mM, the digested nuclei were pelleted and the pellet resuspended and dialysed for 16 h at 5 °C against the same buffer. The supernatant, containing 400–700 A₂₆₀ units of soluble chromatin, was dialysed against 10 mM Tris-HCl pH 8.0, 1.0 mM EDTA and fractionated at room temperature on a Biogel A 150-m column to remove low-molecular weight material. The high-molecular weight fractions were pooled and concentrated by ultrafiltration (Amicon YM10 membrane) to ~10 A₂₆₀ units ml⁻¹, dialysed to 0.35 M NaCl, 20 mM Tris-HCl pH 7.2, 0.2 mM EDTA, stripped of H1 and H5 using Bio-Rad AG 50W-X2 resin¹⁷, dialysed for 16 h at 5 °C against 0.7 M NaCl, 20 mM Tris-HCl pH 7.2, 0.2 mM EDTA to disrupt the native structure, and finally dialysed against the same buffer minus NaCl. In a, 2.0 A₂₆₀ units of this material were digested with 0.1 units of enzyme per µg of DNA at 37 °C for 0, 1, 2, 4 and 8 min (lanes 1–5, respectively), and the DNA was extracted for electrophoresis. Lanes n show DNA fragments produced from a digest of native chromatin (not stripped of H1/H5). b, 'Randomized' chromatin was adjusted to 2.0 mg ml⁻¹ polyglutamic acid (using a neutralized stock solution) and dialysed against 0.1 M NaCl, 20 mM Tris-HCl pH 7.2, 0.2 mM EDTA at a concentration of 2.0 A₂₆₀ units ml⁻¹. Approximately 0.1 vol. of a concentrated solution of salt-extracted H5 in the same buffer containing 2–4 mg ml⁻¹ polyglutamic acid (see ref. 9 for detailed procedures) was added to one portion of a 'randomized' chromatin sample. A blank solution not containing H5 was added to the other portion, and both samples were incubated at 37 °C for 3 h, then dialysed against 20 mM Tris-HCl pH 7.2, 0.2 mM EDTA for digestion. Each sample was digested with 4.0 units of enzyme per µg of DNA at 37 °C for 1, 2, 4 and 8 min; lanes 5–8 correspond to the sample incubated with H5, lanes 1–4 to the sample incubated with polyglutamic acid alone. Lanes n show DNA fragments from digestion of a native chromatin sample; lane s shows DNA fragments produced by a *HincII* digest of RFΦX174 DNA. DNA was extracted and prepared for electrophoresis as described in ref. 9. DNA was electrophoresed on 2.5% agarose gels (40 mM Tris-acetate, 0.5 mM EDTA, pH 8.0), and gels were stained with ethidium bromide and photographed under UV illumination. Polyglutamic acid inhibits micrococcal nuclease, but has very little effect on the appearance of DNA fragments produced from digestion of chromatin⁹. The excess H5 and polyglutamic acid could be removed by pelleting an incubated sample through 5% sucrose. Digestion of this material also produced native-like DNA fragments very similar to those in b, lanes 5–8 (not shown).



When H1/H5-depleted chicken erythrocyte chromatin is incubated for 16 h at 5 °C in a buffer containing 0.7 M NaCl, the regular 212-bp⁶ (average value) nucleosome spacing is disrupted. This is evidenced by the appearance, on digestion with micrococcal nuclease, of DNA fragments of nonspecific lengths and bands that are multiples of ~150 bp, corresponding to closely packed nucleosomes (Fig. 1a). Such rearranged or 'randomized' chromatin is largely a consequence of nucleosome 'sliding'^{7,8}. The digests in Fig. 1a are very similar to those reported for H1-depleted rat liver chromatin incubated for 45 min at 37 °C in 0.6 M NaCl⁷. Moreover, the digests are consistent with the appearance of the rearranged rat liver chromatin in electron micrographs which revealed clusters of tightly packed beads connected by stretches of naked DNA⁷. We emphasize that the chromatin rearrangement is clearly a consequence of the high salt incubation. No rearrangement occurs for carefully H1/H5-stripped chromatin that is maintained at low ionic strength⁷⁻⁹ (not shown).

We have found that when this same preparation of 'randomized chromatin' (Fig. 1a) is incubated with a 10-fold molar excess (over the stoichiometric value) of histone H5 in a buffer containing 0.1 M NaCl and 2 mg ml⁻¹ polyglutamic acid, a

highly ordered and native-like nucleosome arrangement is restored (Fig. 1b, lanes 5–8). The native-like spacing is not restored on incubation with polyglutamic acid in the absence of H5 (Fig. 1b, lanes 1–4), and lesser amounts of H5 are less effective (not shown). Also, H5 in the absence of polyglutamic acid is ineffective (not shown). This suggests that polyglutamic acid may provide a favourable environment for an H5-driven spacing reaction to occur. The band positions in lanes 5 and 6 of Fig. 1b are clearly very similar to those of native chicken erythrocyte chromatin (lanes n), with bands that are multiples of ~200 bp. In addition, the backgrounds and the decrease in length of bands with increasing digestion do not differ significantly from those of a native chicken erythrocyte chromatin sample (not shown).

To remove all traces of the native structure at the start, we incubated a 'randomized chromatin' sample in 0.7 M NaCl for an additional 2 h at 37 °C. Figure 2a shows that this treatment obliterated the 200-bp periodicity even at very early digestion times, which strongly favour detection of longer periodicities. These digests exhibit only 150-bp periodicities superimposed on very high backgrounds. When this extensively disordered chromatin sample was incubated with H5 and polyglutamic

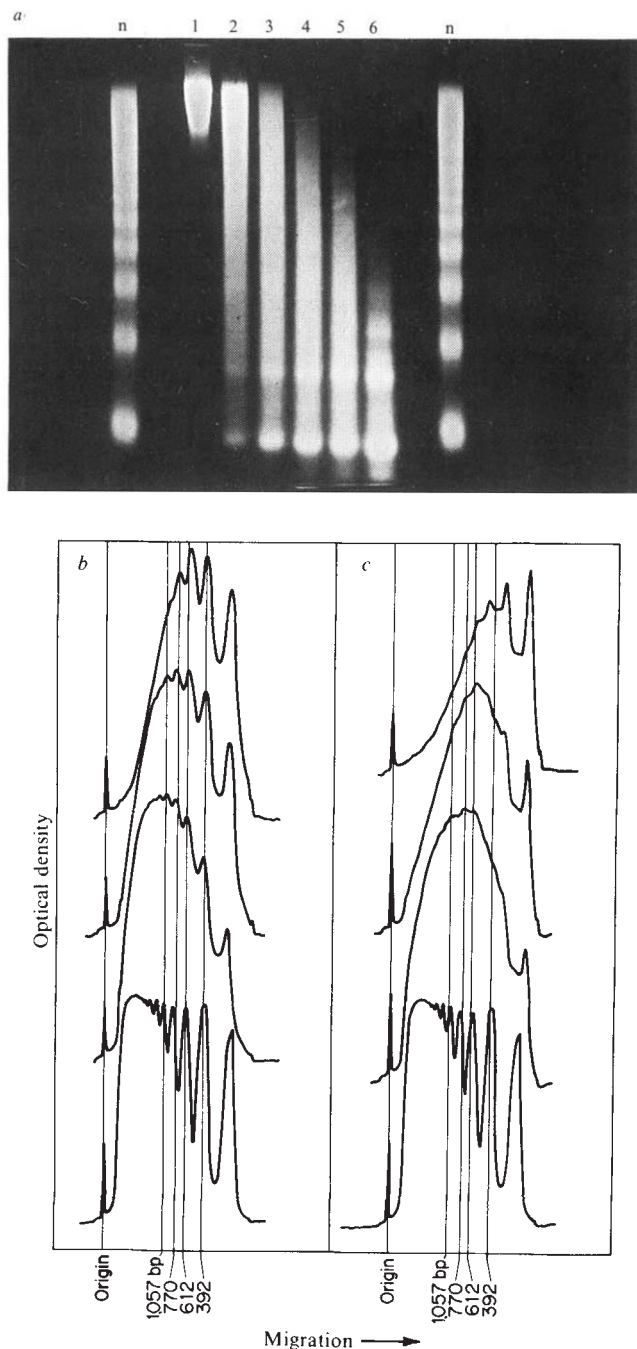


Fig. 2 Electrophoresis of DNA fragments produced by micrococcal nuclease digestion of an extensively rearranged H1/H5-stripped chromatin sample, untreated (a), incubated with H5 and polyglutamic acid (b), or incubated with polyglutamic acid alone (c). Samples were treated as described in Fig. 1 legend. Sample a was digested with 0.04 units of enzyme per μg of DNA at 37°C for 0, 0.5, 1, 1.5, 2 and 4 min (lanes 1–6, respectively). Lanes n show DNA fragments from a digest of native chromatin. Samples b and c were each digested with 2.0 units of enzyme per μg of DNA at 37°C for 1, 2 and 4 min. The three digests are displayed in order of decreasing extent of digestion from top to bottom; the bottom tracing for each was produced from a native chromatin sample. All DNA samples in b and c were electrophoresed on a single gel together with a *HincII* digest of RF Φ X174 DNA for calibration. The photographic negative of the gel was scanned using a densitometer (E-C Apparatus).

native-like bands were observed on the gel (Fig. 2c). In disordered chromatin samples prepared using NaCl concentrations $>0.7\text{ M}$, where histone octamers can dissociate from the DNA¹⁰, or in chromatin reconstituted from core histones and DNA at physiological histone to DNA ratios, only partial restoration of the native nucleosome arrangement occurred (not shown). We have, however, succeeded in producing fairly well resolved native-like nucleosome dimers and trimers in completely reconstituted chromatin samples using core histone to DNA ratios of <0.6 (wt/wt)⁹.

Our results suggest that H5, in the presence of polyglutamic acid, can efficiently 'slide' apart histone octamers in small closely packed clusters of nucleosomes to produce nucleosomes with mature physiological spacings, provided that regions of naked DNA are available for the clusters to fully expand. For DNA molecules or regions that are too densely packed with octamers, as would tend to occur in a portion of a sample reconstituted at physiological histone to DNA ratios, octamer crowding effects would be expected to preclude complete restoration of the native nucleosome periodicity. This is consistent with our observations. Also in accord with this hypothesis, histone octamers in chromatin stripped of H1/H5, but possessing the native internucleosome spacing, do not dissociate from DNA on incubation in a solution containing 0.1 M NaCl and 2 mg ml^{-1} polyglutamic acid for 3 h at 37°C . In fact, these regularly spaced octamers do not even slide together to an appreciable extent in these conditions, as measured by extensive micrococcal nuclease digestion (data not shown). Thus, the energy for sliding octamers apart in the spacing reaction seems to be provided by the interaction of H5 with nucleosomes^{11,12}.

This mechanism is consistent with the observations made in *in vivo* systems^{3–5} and intact nuclei². Our data, before and after incubation with H5, appear very similar to those from digests of newly replicated chromatin undergoing maturation^{2,3,5}. In addition, Laskey and Earnshaw¹³ have suggested that a mechanism must exist for correctly aligning randomly deposited nucleosomes, as assembly in cell-free systems^{14,15} is non-cooperative and, presumably, random. Our data suggest that this mechanism may be the H1-induced sliding apart of octamers in nucleosomes randomly deposited on the DNA.

Our results show, for the first time, that H5 (and presumably H1) is actively involved in nucleosome spacing and that histone modifications are not necessarily required for the spacing reaction to proceed. Also, our *in vitro* system should be well suited for addressing two fundamental problems in chromatin structure: (1) the origin of the tissue- and species-specificity of the nucleosome repeat length, and (2) the cause of the apparent heterogeneity in nucleosome spacing in chromatin¹⁶.

We thank Tamera Townsend and Ruth Gyure for technical assistance. This work was supported by grant GM 27719 (to A.S.) from the USPHS.

Received 5 October 1982; accepted 13 January 1983.

- Kornberg, R. D. *A. Rev. Biochem.* **46**, 931–954 (1977).
- Seale, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2717–2721 (1978).
- Levy, A. & Jakob, K. M. *Cell* **14**, 259–267 (1978).
- Murphy, R. E., Wallace, R. B. & Bonner, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3336–3340 (1980).
- Jackson, V., Marshall, S. & Chalkley, R. *Nucleic Acids Res.* **9**, 4563–4581 (1981).
- Morris, N. R. *Cell* **9**, 627–632 (1976).
- Spadafora, C., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **100**, 225–235 (1979).
- Weischet, W. O. *Nucleic Acids Res.* **7**, 291–304 (1979).
- Künzler, P. & Stein, A. *Biochemistry* (in the press).
- Germond, J. E., Bellard, M., Oudet, P. & Chambon, P. *Nucleic Acids Res.* **3**, 3173–3192 (1976).
- Simpson, R. T. *Biochemistry* **17**, 5524–5531 (1978).
- Thoma, F., Koller, T. H. & Klug, A. *J. Cell Biol.* **83**, 403–427 (1979).
- Laskey, R. A. & Earnshaw, W. C. *Nature* **286**, 763–767 (1980).
- Laskey, R. A., Mills, A. D. & Morris, N. R. *Cell* **10**, 237–243 (1977).
- Nelson, T., Hsieh, T.-S. & Brutlag, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5510–5514 (1979).
- Prunell, A. & Kornberg, R. D. *J. molec. Biol.* **154**, 515–523 (1982).
- Ruiz-Carrillo, A., Puigdomenech, P., Eder, G. & Lurz, R. *Biochemistry* **19**, 2544–2554 (1980).

Provoking thoughts on agriculture

J.C. Bowman

Agricultural Research Policy. By Vernon W. Ruttan.
University of Minnesota Press: 1982. Pp.369. Hbk \$32.50; pbk \$13.95.

In an admirable introduction Professor Ruttan presents his credentials for examining agricultural research policy. For 30 years he has been employed as an economist, mostly in the United States but also in the Philippines and Japan. In recent years the emphasis of his interests has, he says, moved from economics to agricultural research policy, about which he has been a prolific writer.

Professor Ruttan's introduction concludes:

In the book judgements are made about agricultural research policy. These judgements are not confined by any formal philosophical system. They arise out of a combination of scientific analysis, professional experience and my personal perception of responsibility for the results of scientific inquiry and technological development.

He has remained firm to his brief and has not formally reviewed the literature on the subject; indeed there are only a few indications in the text that others may not agree with Professor Ruttan on many of the matters which he discusses. At this stage the reader might be forgiven for some lack of trust in the author's integrity.

Each of the chapters is usually based on one of Professor Ruttan's previous publications or lectures. Whilst the writing is lucid the integration of source material might have been more thorough, reducing the impression that each chapter is an unrelated aspect of a common theme. Another, more serious, criticism is that the distinction between Professor Ruttan's idea of agricultural research policy in general is not always clearly dissociated from his views about American practices. The suspicion lingers until near the end of the book that the author harbours some hope of finding a single optimum solution to the determination and management of agricultural research policy, and that this is not far removed from the current approach in the United States. The last two chapters, however, happily prove such suspicion to be unfounded and make clear the author's continuing pursuit of alternative answers in an international context.

The main purposes of agricultural research might be said to be four-fold: to improve the return on resources used in existing methods of production, processing, distribution, marketing and consumption; to make the final product more acceptable to the consumer in quality and price; to make the existing methods of production etc. more acceptable on social and moral grounds; and to facilitate change in existing processes to meet future needs and opportunities. Such a statement does not appear in Professor Ruttan's

book. Rather he takes these objectives for granted and discusses the achievements of agricultural research in increasing agricultural output per unit of individual and total resource inputs. This is a familiar economic approach. He then tackles some of the factors determining the efficiency of agricultural research, such as the nature of the research institution, national and international systems (with a rather brief reference to Europe and the role of FAO), its funding, and its location and scale.

Professor Ruttan's approach to these topics is somewhat mechanistic. Whilst it is clear that he appreciates the importance of the context in which research is conducted, the effect of the motives of individual scientists on the pertinence and success of a research programme is largely ignored. The desire of research workers to pursue knowledge for its own sake may not always coincide with the objectives held by those who invest in research, and here it would have been interesting to have the author's thoughts on the best means of selecting research managers and scientists, and how best to motivate them.

In considering the respective merits of public and private funding and responsibility, Professor Ruttan sets down the criteria, as he sees them, for the appropriate roles of each in agricultural research. The public sector has been justified, he says, because "in many areas incentives for private sector research have not been adequate to induce an optimum level of research investment", because "its complementarity with education" and because "it has contributed to the maintenance or enhancement of a competitive structure in the agricultural production, input, and marketing sectors". Where, however, are the arguments that public sector research is needed to safeguard society's interests and to undertake basic research for which the private sector will rarely pay, however much money it has to invest? Did Professor Ruttan omit them because he did not consider them important and because he did not consider them valid within his perceived scheme for an improved agricultural research system?

Another interesting topic discussed at length by Professor Ruttan is the selection of a research portfolio. The peer review system, cost-benefit analysis as well as more sophisticated systems for choosing a research programme are considered. But where are the deficiencies and long-term biases of the peer review method, how can the allocation of funds be determined for basic research, when by definition the

results of the research are not known, and how should the appropriateness of research be judged?

For the reader studying agricultural research policy for the first time this book would not be a good starting point because of its selective, at times self-indulgent, air. Conversely, those familiar with the topic will find the book both provocative and stimulating. In the end even the author is left asking questions about the role of social science in agricultural research, about the responsibility to society of research scientists and about the part non-scientists may play in determining the amount and type of research undertaken. It is a measure of the thought-provoking nature of Professor Ruttan's book that one would like to discuss these issues with him. □

J.C. Bowman, formerly Director of the Centre for Agricultural Strategy at Reading University, is Secretary to the Natural Environment Research Council.

From fantasy to fact

Fred Rosen

Monoclonal Antibodies in Clinical Medicine.

Edited by Andrew J. McMichael and John W. Fabre.

Academic: 1982. Pp.661. £34, \$63

IN THEIR paper submitted to *Nature* on 14 May 1975 Köhler and Milstein concluded: "Such cells can be grown *in vitro* in massive cultures to provide specific antibody. Such cultures could be valuable for medical and industrial use" (*Nature* 256, 495; 1975). That may prove to be the scientific understatement of the century. Hybridoma technology has fructified to such an extent as to have become one of the keystones of the biological revolution in progress.

McMichael and Fabre have assembled a notable group of contributors to a volume that lucidly shows what hybridomas have done for medicine and biology. This book is a salubrious mixture of fact and fantasy from its introduction by Milstein, who describes the simple and beautiful experiments that produced this wonder, to the rich speculations of H.S. Kaplan *et al.* on what life might be like if only we had a thesaurus of human hybridomas.

As one proceeds through the plethora of contributions to immunology, haematology, genetics, virology, parasitology, bacteriology, neurobiology and oncology, each of which is supported by graphs, tables and photomicrographs, the reader comes to ponder: "Why monoclonal antibodies? Why not goat or rabbit antibodies?" Near the end of this volume the answer is brilliantly presented in a chapter by E. Haber. With the seemingly trivial aim of discussing monoclonal antibodies to drugs such as digoxin or biologically active

peptides such as angiotensin, Haber launches into a scholarly and wonderful discussion of hybridoma technology as a triumph of logic over empiricism. His argument is echoed elsewhere in the book by P. Goodfellow and Ellen Solomon, who describe with great clarity how monoclonal antibodies were critical to the localization of the human gene for the third component of complement to chromosome 19, and by S. Cohen who recounts how monoclonal antibodies can effectively interrupt the successful life cycle of malarial parasites.

At the end of the book there are four practical, useful chapters on how to make monoclonal antibodies, purify them and use them in immunohistology or in the fluorescence-activated cell sorter. The almost simultaneous appearance of cell sorting and hybridoma technologies greatly facilitated the application of the latter and this is well described in a chapter by P.C.L. Beverly.

Finally an appendix contains a list of monoclonal antibodies that were available last year; the list must be doubled in length at least since this book went to press. McMichael and Fabre have produced a volume that is indispensable to biomedical researchers, and have set themselves a high standard for a repeat performance in the near future. □

Fred S. Rosen is the James L. Gamble Professor of Pediatrics at Harvard Medical School.

ADVERTISEMENT

NUCLEAR ENVIRONMENT

YOUR THING?

Then turn to the #1 source for publications on nuclear energy and related topics

American Nuclear Society

News and developments, journals, critical reviews, transactions, monographs, textbooks, proceedings, public information . . .

**Write today
for your free 108-page
ANS publication catalog.
Do your thing the right way.**

**American Nuclear Society
555 N. Kensington Avenue
La Grange Park, IL 60525 USA
312/352-6611**

Some novel truths of medicine

Douglas Black

Magic Shots: A Human and Scientific Account of the Long and Continuing Struggle to Eradicate Infectious Diseases by Vaccination.

By Allan Chase.

Morrow: 1982. Pp. 576. \$17.95.

"A HUMAN and scientific account of the long and continuing struggle to eradicate infectious diseases by vaccination". On seeing that ominous phrase on the front of the dust jacket, I expected the worst of this book — science and medicine made trivial through the guise of a fictionalized documentary. Only on reading it did I discover that the blemishes of fiction are on the surface, whereas the substance of the book is thorough and in some respects potentially valuable.

Before turning to the merits of the book, and its possible value to readers, let me justify by illustration my first rather negative reaction. On p.28 we read, "Casper was not terribly surprised after this to be reassigned from Chicago to a gonorrhoea center in Jackson, Mississippi". And on p.182 we find, "Park was flabbergasted". Such insubstantial grace-notes, however plausible, must surely be conjectural; far from enhancing the interest of the narrative, they marginally diminish both its credibility and the pleasure of reading it.

I must now say that on getting into the book I found it a most worthy piece of popular science. Of course the author has a splendid story to tell, from the early days of Jenner to the possibilities being opened up by current advances in immunology. The historical sections, which make up the greater part of the book, have been carefully researched and are adequately referenced. The information provided on immunology is up to date, including awareness of the potential of monoclonal antibodies in the treatment of cancer and of malaria. This would not, though, be a book which would be bought by microbiologists or immunologists for its scientific content; but such readers could well find interest in its historical sections. As a general physician, I learned much on both these aspects. The most important readership I judge to be twofold — the young who are contemplating a medical or biological career, and who may have become infected by the notion that all the great discoveries have already been made; and the intelligent lay public, who may be culpably unaware that medicine, in both its preventive and its curative aspects, is at its most effective when it is science-based. To be sure skills other than the scientific are still required to ensure the application of

scientific knowledge, whether to the individual or to the population; but the most necessary ingredient of successful communication is to have something which is worth communicating. I have taken quite a number of words to say what Joseph L. Melnick says in a few on the cover blurb — "*Magic Shots* reads like a novel, but it is all true".

In his perspective of infectious diseases, Chase notes the importance of social factors, but also recognizes the importance from the mid-twentieth century onwards of the treatment of infection in the individual. He nonetheless rightly challenges the view that treatment with antibiotics can be the complete answer to infective illness. The emergence of resistant organisms, and the increased incidence of serious infection in patients with natural or acquired defective immunity, bring new problems, some of which may find their solution in new vaccines based on advances in genetics and immunochemistry. But he also points out that progress in these matters depends on a continuing supply of scientists, and that it is essential to awaken an interest in science at an early stage of education. His strength of feeling on this matter can be gauged from two quotations — "An education that makes science and all other aspects of our culture the legacy of young people should begin early and must continue through and beyond graduate school. It has to be good, and it has to be available to all children of all socioeconomic classes". And, in relation to adverse effects on education of the doubling of military expenditure in the United States, he says "Education, in our angry and malignantly anti-intellectual times, has become expendable".

The book includes a foreword written by D. A. Henderson, the former Director of the WHO Smallpox Eradication Programme. Henderson not surprisingly calls for further support of vaccine research, but in my view on grounds which could be misleading. What he says is "To the extent that we are willing to commit resources to the development of these powerful weapons to prevent illness, the sooner we shall shift our focus from treating the sick to sustaining the healthy". The supporting clause seems to me to misconceive the totality of the medical task. Of course, basic and applied work on vaccines must go on, and increase; but not at the expense of curing when possible, and at the very least caring for those with non-infective diseases. The title of the first chapter in the book, "Prevention is Better", begs the question — better than what? Basic medical science should contribute to both prevention and treatment, and both are essential to medical work. Prevention is important for the healthy, treatment for those who are already ill. □

Sir Douglas Black is President of the Royal College of Physicians and Emeritus Professor of Medicine at the University of Manchester.

Contrast of resource frontiers

G. de Q. Robin

Arctic and Antarctic: A Modern Geographical Synthesis.

By David Sugden.

Basil Blackwell/Barnes & Noble: 1982.

Pp. 472. £29.50, \$35.

Arctic and Antarctic was a title used previously in 1939 by Colin Bertram with the subtitle *The Technique of Polar Travel*. In that book, published by Cambridge University Press, Bertram discussed basic principles of travel in an age when dog teams had not given way entirely to tractors and aeroplanes. Forty years later, David Sugden writes with an extensive but different experience of an Arctic and Antarctic that are no longer so remote. Commerce, armed services, scientists and governments have become increasingly active in these "resource frontier regions" of geographers.

The new volume is sub-titled *A Modern Geographical Synthesis*. It is written to introduce undergraduates reading geography to the polar regions and to use this knowledge to test geographical concepts. The coverage is wide rather than deep in both time and space, as recognized by the author who has aimed at geographical completeness.

Sugden begins by placing the natural systems of the Arctic and Antarctic in context, with chapters on plate tectonics and climate, which are followed by regional descriptions of the glacial system, the periglacial system and the polar marine system. Photographs and diagrams as well as his own wide experience help the author to cover this broad field in a mere 140 pages. Then follows a general survey of present knowledge of environmental change in polar regions.

The second section, dealing with human systems, starts with a concise but perceptive outline of the exploration, settlement and development of polar regions. After an account of the indigenous population of the Arctic, 35 pages or so are devoted to a survey of each of five regions: Greenland and Svalbard, Arctic Canada, Alaska, the Soviet Union and Antarctica. Each starts with around four pages on "Physical constraints and resources" followed by a few pages on "The human spatial pattern", "The evolution of the spatial system", then some discussion of "The functioning of the spatial system" such as transport networks. Special topics such as comparison of Antarctic with Arctic developments follow before "Conclusions" about each region.

The geographical systems approach provides a useful link between the pattern of development in the Arctic and those in

other less developed regions such as Africa. However, in Antarctica, this seems somewhat forced as scientific knowledge is presented as yet another resource. The stimulus from science in opening up the Antarctic continent is in a different class to the economic factors operating in other polar regions, and one regrets this was not given more emphasis.

The book is well written and illustrated and follows a logical pattern. This makes it useful for those not familiar with the different regions discussed. Inevitably, however, some detailed points of the author's presentation could be brought up to date or improved. For example, the large proportion of frazil ice found in many

drifting ice floes is not mentioned. Elsewhere Sugden says that the Weddell Sea drift was discovered by Ernest Shackleton when *HMS Endurance* (sic) was frozen into the Weddell Sea in 1915 — although this had been experienced by Filchner in the *Deutschland* in 1911–1912 as Shackleton will have known before setting out on *S. Y. Endurance*. These are however blemishes of detail on a broad survey that benefits, at least in the physical sciences, from drawing on contrasts between conditions in the Arctic and Antarctic. □

G. de Q. Robin was formerly Director of the Scott Polar Research Institute, Cambridge.

William Fowler and the elements

M. G. Edmunds

Essays in Nuclear Astrophysics.

Edited by C.A. Barnes, D.D. Clayton and D.N. Schramm.

Cambridge University Press: 1982.

Pp. 560. Hbk £40, \$75; pbk \$29.95.

ORIGINS have their own peculiar fascination, and the origin of a particularly basic part of us — the chemical elements — is no exception. The realization that nuclear fusion was the only source of energy sufficient to power the stars led inevitably to the idea that further reactions could build up complex atomic nuclei from hydrogen and helium.

It is now over 25 years since the immensely fruitful collaboration between nuclear physics and astronomy was made apparent by the publication in 1957 of the classic systematizations of the major element production processes by Burbidge, Burbidge, Fowler and Hoyle, and by Cameron. The details of the processes have subsequently been elaborated and changed, but the broad outline remains remarkably the same. Professor Willy Fowler of the California Institute of Technology has been a prime force behind studies of nuclear astrophysics since the late 1940s. This wide-ranging collection of essays by his present and former colleagues and students is a testament to his influence and enthusiasm, particularly since almost every topic covered has been worked on or encouraged by him.

There are presently believed to be three major sites for the synthesis of the elements — the early Universe, where extreme temperatures and densities produce helium and a few light elements; the interstellar medium, where cosmic ray particles interact with interstellar atoms to produce a few more light elements; and nuclear reactions in stars, where elements including and heavier than carbon are produced, particularly during violent explosive

evolution. One of the principal problems in nuclear astrophysics is the determination of the rates of reactions which occur during the normal quiescent life of a star, because the energies at which the reactions occur are such that laboratory simulations of the same conditions would give unmeasurably slow rates. Several of these essays address the problem of extrapolating laboratory measurements to the relevant energies, important for astrophysicists in predicting the composition of a star after its quiescent life, and a complicated problem in nuclear physics in its own right. The major skeleton in the stellar evolution closet has an excellent airing in an anecdotal article on the solar neutrino problem. The very low (compared to theoretical estimates) flux of observed neutrinos from one of the side chains in the solar core nuclear reactions still persists, although the impending construction of a new experiment based on a gallium detector to look at neutrinos from the main proton–proton reaction should resolve whether the astrophysicists have made their stellar models wrong or whether neutrinos do not have the properties that are usually assumed. Either way, astrophysics or particle physics will benefit from the result.

The nuclear physics of exploding stars is in some ways easier, since the reaction energies come into the range accessible in laboratory equipment. But the extreme physical conditions and hydrodynamics make the theoretical models of stars much more uncertain. It seems likely, as discussed in several of these essays, that the generation of a shock wave in the final stages of evolution of massive stars must provide sudden high temperatures for nucleosynthetic reactions and also give sufficient push to expel some of the products from the collapsing star. The details of this complex process remain uncertain, as they do for the other main types of explosion involving lower-mass, possibly binary, stars.

In recent years much excitement has come as a result of beautiful analytic chemistry which has discovered marked anomalies in relative isotope abundances in certain meteorites. Since this material is

believed to represent a sample of the very early Solar System, these anomalies are often interpreted as the result of an injection of freshly-synthesized material into the proto-planetary nebula just before, or while, it formed. Any glimpse of conditions in this elusive era is welcome, and fairly detailed review of some of the short-lived nuclides appears in this volume.

The slow and rapid neutron capture chains synthesize the heavy elements beyond iron in the periodic table, but the actual sites where they occur remain rather uncertain. That the slow addition of neutrons occurs in evolved red giant stars is witnessed by the subsequent chemical peculiarity of their atmospheres, and the fast addition probably occurs near the cores of supernovae explosion of stars. It is still difficult to be sure of what reactions are most important in supplying the free neutrons, or of the conditions under which addition occurs. Indeed, it is argued in one of these essays that the rapid process is not really as fast as previously believed.

Nuclear astrophysics is an active field. This is a valuable book for those who know

the basics of nuclear astrophysics and who want to find out about many of the current problems, but it is obviously not intended to be at a level suitable for a complete beginner. Many research workers will find it an essential source for Cameron's latest version (the last was in 1973) of his compilation of measured abundances in the Solar System. The book's topics are many, including hopes for detection of γ -rays from nucleosynthetic reactions and brief discussion of the early Universe, in addition to everything mentioned above. There is not a great deal included on fitting the nucleosynthesis processes into the overall evolution of galaxies, but to treat this fully would probably have required another volume, and the limitation is sensible.

Shining through the personal remarks in these essays is Willy Fowler's inspiring personality, and his practised adage that "Physics is fun!". □

M.G. Edmunds is a Lecturer in the Department of Applied Mathematics and Astronomy, University College, Cardiff.

Vision from the womb

F.W. Campbell

A Biography of the Eye: Development, Growth, Age.

By R.A. Weale.

H.K. Lewis: 1982. Pp.366. £17.

PROFESSOR Weale introduces his book with "Some twenty years ago I wrote a survey of ocular aging [*The Aging Eye*, H.K. Lewis; 1963]. It aroused more interest than one would normally hope for with such a specialised study, and I have been urged ever since to write a new edition. The sheer boredom of reheating a meal put me off".

We should be grateful that he made this decision, for his new book embraces the changes that occur in the visual process from prenatal development until blindness and death. His Fig.1.1 typifies his project and teaching style where he plots percentage survival data for cafeteria tumbler, comparing annealed and toughened glass. He proceeds lucidly and succinctly to describe the reasons for the increasing age of the population. As he puts it: "Mathematically speaking, it is permissible to say that those who prolong life promote blindness".

In pursuing his theme, Weale has uncovered a large number of relevant and interesting references which will be unknown to the great majority of scientists. He blends them into a coherent and interesting account of the problem of blindness in the future which will encourage much-needed research in this area. The book quotes

many new techniques for investigating visual performance not only in infants but also in adult blindness. The variety and quality of the figures and half-tone plates is superb. Many teachers will find them extremely useful in illustrating their lecture courses.

Recently there has been a plethora of books on vision, most of them saying the same thing. By contrast Professor Weale's book is quite unique and deals with almost everything the other books omit. It will also have a strong appeal to those non-visual scientists and clinicians interested in the aging process and eye development. □

F.W. Campbell is Reader in Neurosensory Physiology at the University of Cambridge.

Plants and scholars

Peter D. Moore

History of Botanical Science.

By A.G. Morton.

Academic: 1982. Pp.475.

Hbk £19.80, \$45.50; pbk £8.80, \$21.50.

BOTANY, according to A.G. Morton, has its origin in magic and medicine. Since then, the cynic might argue, it has tended to move away from medicine. Its history, at least during the early stages, was quite independent of other sciences, even zoology, perhaps because of its association with agriculture. It is this independence which justifies the writing of a history of botany rather than one dealing with biology in general. Added to which, such a history has not been written in English for some 40 years.

History needs to be rewritten periodically, partly because it is constantly being added to, partly because historical research is constantly revealing new information and evidence concerning the past and partly because viewpoint and perspective are in constant flux as scientific thinking evolves. Our present view of Aristotle for example, may differ from the way Julius von Sachs regarded him.

Morton has not dwelt upon current developments in botany. In fact only 19 pages out of a total of 466 are devoted to the twentieth century, so one must look for other areas in which he has added to historical knowledge. One finds a freshness of approach in the weight Morton places upon certain concepts and the work of certain early botanists. Take Theophrastus, for example; Morton emphasizes the value of Theophrastus' concept of the adaptation of plants to their environment and regards him as the father of ecology. In his assessment of the work of the classical botanists of the ancient world, one is struck by the detail and scholarship of Morton's approach. The subtleties of the Greek language are well appreciated and frequently quoted by him.

Although the author makes full use of his knowledge of verbal nuances in the interpretation of developments in botanical concepts, I feel that he could have taken greater advantage of modern developments in botanical science in order to interpret early discoveries and notions. The assessment of the work of Dioscorides, for example, and its influence upon Islamic herbalism, makes a fascinating story which could have been made yet more vivid if it had been considered in the light of modern biochemical knowledge of the plants concerned.

The book develops along strictly chronological and historical lines rather than by topics, and this enables the author to present the development of botany in its contemporaneous intellectual environment. Occasionally, however, one feels that something is lost by separating botany so strictly from zoology; zoological analogies often played an important part in the development of botanical models, a point clearly illustrated by François Delaport in his analysis of the use of zoological analogies in eighteenth-century botany (*Nature's Second Kingdom*, published by MIT Press in 1982).

This book is a work of great care and scholarship and it traces in detail the emergence of philosophical concepts within the science of botany. It is a book which will delight the historian, but may leave the modern botanist just a little disappointed in that the concepts themselves are not always set out in ways which clarify their relevance to current botanical thought. □

Peter D. Moore is Senior Lecturer in the Department of Plant Sciences, King's College, University of London.

Mathematics

- ASHURST, F.G. *Founders of Modern Mathematics*. Pp.145. ISBN 0-584-10380-8. (Muller, London: 1982.) £7.50.
- DE OLIVEIRA, J.T. and EPSTEIN, B. (eds). *Some Recent Advances in Statistics*. Pp.248. ISBN 0-12-691580-6. (Academic: 1982.) £24, \$49.50.
- LIONS, J.L. *Some Methods in the Mathematical Analysis of Systems and Their Control*. Pp.542. ISBN 0-677-60200-6. (Science Press: 1981.) Np.
- LUKES, D.L. *Differential Equations: Classical to Controlled*. Pp.322. ISBN 0-12-459980-X. (Academic: 1982.) \$37.50.
- OVSIIANNIKOV, L.V. *Group Analysis of Differential Equations*. Pp.416. ISBN 0-12-531801-1. (Academic: 1982.) \$54.
- WALTER, E. *Identifiability of State Space Models: With Applications to Transformation Systems*. Lecture Notes in Biomathematics, Vol.46. Pp.202. Pbk ISBN 3-540-11590-0/0-387-11590-0. (Springer-Verlag: 1982.) DM 32.50, \$13.60.
- WHITEMAN, J.R. (ed.). *The Mathematics of Finite Elements and Applications IV*. Pp.555. ISBN 0-12-747254-1. (Academic: 1982.) £22, \$40.50.

Astronomy

- Astronomy and Astrophysics for the 1980s*, Vol.1. Report of the Astronomy Survey Committee. Pp.189. Pbk ISBN 0-309-03249-0. (National Academy Press: 1982.) £13.25, \$23.10.
- BURBIDGE, G. *et al.* (eds). *Annual Review of Astronomy and Astrophysics*, Vol.20. Pp.631. ISBN 0-8243-0920-0. (Annual Reviews Inc: 1982.) Np.
- HER MAJESTY'S STATIONERY OFFICE. *The Astronomical Almanac for the Year 1983*. ISBN 0-11-886914-0. (HMSO: 1982.) Np.

Physics

- BEHRENS, H. and BUHRING, W. *Electron Radial Wave Functions and Nuclear Beta-Decay*. Pp.626. ISBN 0-19-851297-X. (Oxford University Press: 1982.) £55.
- BOWLER, K.C. and SUTHERLAND, D.G. (eds). *Gauge Theories and Experiments at High Energies*. Proceedings of the 21st Scottish Universities Summer School in Physics held August 1980, St. Andrews. Pp.640. ISBN 0-905945-04-2. (Scottish Universities Summer School in Physics: 1981.) £20.
- CAIRNS, R.A. and SANDERSON, J.J. (eds). *Laser-Plasma Interactions*. Proceedings of the 20th Scottish Universities Summer School in Physics held August 1979, St. Andrews. Pp.840. ISBN 0-905945-03-4. (Scottish Universities Summer School in Physics: 1980.) £18.
- DARROZES, J.S. and FRANÇOIS, C. *Mécanique des Fluides Incompressibles*. Lecture Notes in Physics, Vol.163. Pp.459. Pbk ISBN 3-540-11578-1/0-387-11578-1. (Springer-Verlag: 1982.) DM 45, \$18.80.
- DEMOKAN, M.S. *Mode-Locking in Solid-State and Semiconductor Lasers*. Pp.227. ISBN 0-471-10498-1. (Wiley: 1982.) £18.50, \$39.95.
- FIECK, G. *Symmetry of Polycentric Systems*. Lecture Notes in Physics, Vol.167. Pp.137. Pbk ISBN 3-540-11589-7/0-387-11589-7. (Springer-Verlag: 1982.) DM 18, \$7.60.
- FISCHER, Ø. and MAPLE, M.B. (eds). *Superconductivity in Ternary Compounds I: Structural, Electronic, and Lattice Properties*. Topics in Current Physics, Vol.32. Pp.283. ISBN 3-540-11670-2/0-387-11670-2. (Springer-Verlag: 1982.) DM 75, \$31.30.
- GAY, P. *An Introduction to Crystal Optics*. Pp.262. Pbk ISBN 0-582-30112-2. (Longman: 1982.) £6.95.

Chemistry

- TRAHANOVSKY, W.S. (ed.). *Oxidation in Organic Chemistry*, Part D. Pp.360. ISBN 0-12-697253-2. (Academic: 1982.) \$74.
- WEST, E.G. *Copper and Its Alloys*. Pp.241. ISBN 0-85312-505-8/0-470-27533-2. (Ellis Horwood/Halsted: 1982.) £25, \$50.25.

Technology

- BAIJAL, M.D. (ed.). *Plastics Polymer Science and Technology*. Pp.945. ISBN 0-471-04044-4. (Wiley: 1982.) £115.
- BRAUN, E. and MACDONALD, S. *Revolution in Miniature: The History and Impact of Semiconductor Electronics Re-explored in an Updated and Revised 2nd Edn*. Pp.247. Hbk ISBN 0-521-24701-2; pbk ISBN 0-521-28903-3. (Cambridge University Press: 1982.) Hbk £15; pbk £5.95.
- DAHLBERG, E.C. *Applied Hydrodynamics in Petroleum Exploration*. Pp.161. ISBN 0-387-90677-0/3-540-90677-0. (Springer-Verlag: 1982.) DM 48, \$21.40.
- KRESTA, J.E. (ed.). *Reaction Injection Molding and Fast Polymerization Reactions*. Polymer Science and Technology, Vol.18. Pp.302. ISBN 0-306-41120-2. (Plenum: 1982.) \$42.50.
- LOVERING, D.G. (ed.). *Molten Salt Technology*. Pp.533. ISBN 0-306-41076-1. (Plenum: 1982.) \$65.
- SHAW, R. *Wave Energy: A Design Challenge*. Pp.202. UK Hbk ISBN 0-85312-382-9; pbk ISBN 0-85312-504-X; US ISBN 0-470-27539-1. (Ellis Horwood/Halsted: 1982.) £18.50, \$37.20.

Computer Science

- CLARK, K.L. and TARNLUND, S.-A. (eds). *Logic Programming*. APIC Studies in Data Processing, No.16. Pp.366. ISBN 0-12-175520-7. (Academic: 1982.) £16.80, \$34.50.
- PRESTON, K. Jr. and UHR, L. (eds). *Multi-computers and Image Processing: Algorithms and Programs*. Pp.470. ISBN 0-12-564480-9. (Academic: 1982.) \$34.
- SCHEUERMANN, P. (ed.). *Improving Database Usability and Responsiveness*. Pp.450. ISBN 0-12-624080-9. (Academic: 1982.) \$34.
- TROPPER, C. *Local Computer Network Technologies*. Pp.144. ISBN 0-12-700850-0. (Academic: 1981.) \$19.50.
- WULFORST, H. *Breakthrough to the Computer Age*. Pp.185. ISBN 0-684-17499-5. (Scribners: 1982.) \$12.95.
- YOELI, P. *Cartographic Drawing with Computers*, Vol.8. Pp.137. (Department of Geography, University of Nottingham: 1982.) £4.
- YOYITS, M.C. (ed.). *Advances in Computers*, Vol.21. Pp.452. ISBN 0-12-012121-2. (Academic: 1982.) \$52.

Earth Sciences

- AVEROUS, P. *Tell Me About — The Earth and The Sky*. Pp.61. ISBN 0-00-195833-X. (Collins: 1982.) £3.50.
- BORRADAILE, G.J., BAYLY, M.B. and POWELL, C.McA. (eds). *Atlas of Deformational and Metamorphic Rock Fabrics*. Pp.551. ISBN 3-540-11278-2/0-387-11278-2. (Springer-Verlag: 1982.) DM 138, \$61.30.
- Carbon Dioxide and Climate: A Second Assessment*. Pp.72. Pbk ISBN 0-309-03285-7. (National Academy Press: 1982.) £6.50, \$11.35.

Biological Sciences

- AIZAWA, H. *Metabolic Maps of Pesticides*. Pp.243. ISBN 0-12-046480-2. (Academic: 1982.) \$35.
- ALLEN, T.F.H. and STARR, T.B. *Hierarchy: Perspectives for Ecological Complexity*. Pp.310. ISBN 0-226-01431-2. (University of Chicago Press: 1982.) \$27.50.
- AMBROSE, E.J. *The Nature and Origin of the Biological World*. Pp.190. Hbk ISBN UK 0-85312-454-X, US 0-470-27513-8; pbk ISBN UK 0-85312-455-8, US 0-470-27514-6. (Ellis Horwood/Halsted: 1982.) Hbk £17.50, \$35.15; pbk np.
- ANDERSON, P. *The Wildlife of Mountains and Moorlands*. Pp.45. ISBN 0-356-07124-3. (Macdonald & Co: 1982.) £3.95.
- ASHBURNER, M., CARSON, H.L. and THOMPSON, J.N. Jr. (eds). *The Genetics and Biology of *Drosophila**, Vol.3b. Pp.428. ISBN 0-12-064946-2. (Academic: 1982.) £47.80, \$98.

- BAKER, J.R. (ed.). *Perspectives in Trypanosomiasis Research*. Proceedings of the 21st Seminar held September 1981, London. Pp.105. ISBN 0-471-10478-7. (Wiley: 1982.) £12.95, \$29.95.
- BERKOVITZ, B.K.B., MOXHAM, B.J. and NEWMAN, H.N. (eds). *The Periodontal Ligament in Health and Disease*. Pp.470. Hbk ISBN 0-08-024412-2; pbk ISBN 0-08-024411-4. (Pergamon: 1982.) Hbk np; pbk £24.50, \$49.
- BERRY, M. (ed.). *Growth and Regeneration of Axons in the Nervous System*. Bibliotheca Anatomica No.23. Selected Papers Presented at the 6th European Anatomical Congress held September 1981, Hamburg. Pp.126. ISBN 3-8055-3528-7. (Karger: 1982.) SwFr. 114, DM 137, \$68.50.
- BLOOM, F., BARCHAS, J. *et al.* *Beta-Carbolines and Tetrahydroisoquinolines*. Progress in Clinical and Biological Research, Vol.90. Proceedings of a Workshop held December 1981, California. Pp.454. ISBN 0-8451-0090-4. (Alan R. Liss: 1982.) £35.60, DM 124.
- BONE, Q. and MARSHALL, N.B. *Biology of Fishes*. Pp.253. UK Pbk ISBN 0-216-91017-X; US ISBN 0-412-00151-9. (Blackie/Chapman & Hall in association with Methuen: 1982.) UK £11.95; US np.
- BORNKAMM, R., LEE, J.A. and SEAWARD, M.R.D. (eds). *Urban Ecology*. The 2nd European Ecological Symposium held September 1980, Berlin. Pp.370. ISBN 0-632-00943-8. (Blackwell Scientific Publications: 1982.) £27.50.
- BURNS, R.G. and SLATER, J.H. (eds). *Experimental Microbial Ecology*. Pp.683. ISBN 0-632-00765-6. (Blackwell Scientific Publications: 1982.) £55.
- BURSTEIN, M. and LEGMANN, P. *Lipoprotein Precipitation*. Monographs on Atherosclerosis, Vol.11. Pp.132. ISBN 3-8055-3512-0. (Karger: 1982.) SwFr. 97, DM 116, \$58.25.
- CLEMENTS, C.D. *Medical Genetics Casebook: A Clinical Introduction to Medical Ethics Systems Theory*. Pp.256. ISBN 0-89603-033-4. (Humana Press: 1982.) \$29.50 (US), \$39.50 (elsewhere).
- DARLEY, W.M. *Algal Biology: A Physiological Approach*. Basic Microbiology, Vol.9. Pp.168. Pbk ISBN 0-632-00608-0. (Blackwell Scientific Publications: 1982.) £7.80.
- DELANY, M.J. *Mammal Ecology*. Tertiary Level Biology. Hbk ISBN 0-216-91310-1; pbk ISBN 0-216-91309-8. (Blackie: 1982.) Hbk £17.50; pbk £8.25.
- DESNIKER, R.J., PATTERSON, D.F. and SCARPELLI, D.G. (eds). *Animal Models of Inherited Metabolic Diseases*. Progress in Clinical and Biological Research, Vol.94. Proceedings of the International Symposium held October 1981, Bethesda. Pp.544. ISBN 0-8451-0094-7. (Alan R. Liss: 1982.) £43.
- DURRELL, G. *The Amateur Naturalist*. Pp.320. ISBN 0-241-10841-1. (Hamish Hamilton: 1982.) £12.50.
- EDNEY, A.T.B. (ed.). *Dog and Cat Nutrition*. A Handbook for Students, Veterinarians, Breeders and Owners. Pp.128. Hbk ISBN 0-08-028891-X; pbk ISBN 0-08-028890-1. (Pergamon: 1982.) Hbk £12, \$24; pbk £6, \$12.
- ETHERINGTON, J.R. *Environment and Plant Ecology*, 2nd Edn. Pp.487. Hbk ISBN 0-471-10136-2; pbk ISBN 0-471-10146-X. (Wiley: 1982.) Hbk £25, \$59.95; pbk £12, \$28.
- FAHN, A. *Plant Anatomy*, 3rd Edn. Pp.544. Hbk ISBN 0-08-028030-7; pbk ISBN 0-08-028029-3. (Pergamon: 1982.) Hbk £30, \$60; pbk £13.50, \$27.
- FRECKMAN, D.W. (ed.). *Nematodes in Soil Ecosystems*. Pp.206. ISBN 0-292-75526-0. (University of Texas Press: 1982.) £15.
- GANTEN, D. and PFAFF, D. (eds). *Adrenal Actions on Brain*. Current Topics in Neuroendocrinology, Vol.2. Pp.153. ISBN 3-540-11126-3/0-387-11126-3. (Springer-Verlag: 1982.) DM 78, \$34.70.
- GOLDBERGER, R.F. and YAMAMOTO, K.R. (eds). *Biological Regulation and Development*, Vol.3A: *Hormone Action*. Pp.346. ISBN 0-306-40925-9. (Plenum: 1982.) \$39.50.
- GOTO, Y., HORIUCHI, A. and KOGURE, K. (eds). *Diabetic Neuropathy*. Proceedings of the International Symposium held September 1981, Tokyo. Pp.389. UK ISBN 90-219-0581-7; US ISBN 0-444-90259-7. (Excerpta Medica/Elsevier Biomedical/North Holland: 1982.) \$91, Dfl.195.

- HALL, J.L., FLOWERS, T.J. and ROBERTS, R.M. *Plant Cell Structure and Metabolism*, 2nd Edn. Pp.543. Pbk ISBN 0-582-44408-X. (Longman: 1982.) £12.95.
- HALLET, F.R., STINSON, R.H. *et al.* *Physics for the Biological Sciences: A Topical Approach to Biophysical Concepts*. Pp.255. UK Pbk ISBN 0-412-24750-X; US ISBN 0-458-95280-X. (Chapman & Hall/Methuen: 1982.) UK £9.95; US np.
- HARRINGTON, F.H. and PAQUET, P.C. (eds). *Wolves of the World: Perspectives of Behavior, Ecology, and Conservation*. Pp.474. ISBN 0-8155-0905-7. (Noyes: 1982.) \$48.
- HASKILL, S. (ed.). *Tumor Immunity in Prognosis: The Role of Mononuclear Cell Infiltration*. Immunology Series, Vol.18. Pp.448. ISBN 0-8247-1837-2. (Dekker: 1982.) SwFr.185.
- HAZARD, E.B. *The Mammals of Minnesota*. Pp.280. Hbk ISBN 0-8166-0949-7; pbk ISBN 0-8166-0952-7. (University of Minnesota: 1982.) Hbk \$39.50; pbk \$15.95.
- HOCH, Z. and LIEF, H.I. (eds). *Sexology: Sexual Biology, Behavior and Therapy*. Selected Papers of the 5th World Congress held June 1981, Israel. Pp.422. UK ISBN 90-219-0556-3; US ISBN 0-444-90260-0. (Excerpta Medica/Elsevier Biomedical/North Holland: 1982.) Dfl.230, \$107.
- HUTSON, D.H. and ROBERTS, T.R. (eds). *Progress in Pesticide Biochemistry*, Vol.2. Pp.226. ISBN 0-471-10118-4. (Wiley: 1982.) £22.50, \$49.50.
- HOUSLAY, M.D. and STANLEY, K.K. *Dynamics of Biological Membranes: Influence on Synthesis, Structure and Function*. Pp.330. Hbk ISBN 0-471-10080-3; pbk ISBN 0-471-10095-1. (Wiley: 1982.) Hbk £23, \$52.65; pbk ISBN £8.90, \$19.95.
- JACKMAN, B. and SCOTT, J. *The Marsh Lions*. Pp.224. ISBN 0-241-10827-6. (Elm Tree Books: 1982.) £12.50.
- JEFFCOATE, S.L. (ed.). *Androgens and Anti-Androgen Therapy. Current Topics in Reproductive Endocrinology*, Vol. 1. Pp.188. ISBN 0-471-10154-0. (Wiley: 1982.) £18.50.
- JONES, R.N. and REES, H.B. *Chromosomes*. Pp.254. ISBN 0-12-390060-3. (Academic: 1982.) £24, \$44.50.
- JOSEPH, J. *A Textbook of Regional Anatomy*. Pp.519. Hbk ISBN 0-333-28911-0; pbk ISBN 0-333-28912-9. (Macmillan Press, London: 1982.) Hbk £30; pbk £15.
- KIRCHNER, H. *Immunobiology of Infection with Herpes Simplex Virus*. Monographs in Virology, Vol. 13. Pp.104. ISBN 3-8055-3517-1. (Karger: 1982.) SwFr. 65, DM 78, \$39.
- KLINGMAN, G.C., ASHTON, F.M. and NOORDHOFF, L.J. *Weed Science: Principles and Practices*, 2nd Edn. Pp.449. ISBN 0-471-08487-5. (Wiley: 1982.) £19, \$32.50.
- LAURSEN, G.A. and AMMIRATI, J.F. (eds). *Arctic and Alpine Mycology. The First International Symposium*. Pp.559. ISBN 0-295-95856-1. (University of Washington Press: 1982.) £31.50.
- LUDERER, A.A. and WEETALL, H.H. (eds). *Clinical Cellular Immunology: Molecular and Therapeutic Reviews*. Pp.416. ISBN 0-89603-011-3. (Humana: 1982.) \$49.50 (US), \$59.50 (elsewhere).
- MACLEAN, W.P. *Reptiles and Amphibians of the Virgin Islands*. Pp.54. Pbk ISBN 0-333-30363-3. (Macmillan Education: 1982.) £3.25.
- MAKITA, A., HANDA, S. *et al.* (eds). *New Vistas in Glycolipid Research*. Pp.490. ISBN 0-306-41108-3. (Plenum: 1982.) \$62.50.
- MARSHALL, N.B. *Developments in Deep-Sea Biology*. Pp.566. ISBN 0-7137-0797-6. (Blandford Press: 1982.) £8.95.
- MATHIAS, M.E. (ed.). *Flowering Plants in the Landscape*. Pp.254. ISBN 0-520-04350-2. (University of California Press: 1982.) £12.75, \$22.
- MATTISON, C. *The Care of Reptiles and Amphibians in Captivity*. Pp.304. ISBN 0-7137-1158-2. (Blandford: 1982.) £8.95, \$17.95 (US), \$22.95 (Canada).
- McKAY, F.W., HANSON, M.R. and MILLER, R.W. *Cancer Mortality in the United States: 1950-1977*. Pp.475. NIH Publication No.82-2435. (US Department of Health and Human Services: 1982.) \$18 (US, Canada, Mexico), \$22.50 (elsewhere).
- MILLER, J. and VAN LOON, B. *Darwin for Beginners*. Pp.175. Hbk ISBN 0-906495-95-4; pbk ISBN 0-906495-96-2. (Writers and Readers: 1982.) Hbk \$12.95, £7.95; pbk np.
- MOYNIHAN, M. and RODANICHE, A.F. *The Behaviour and Natural History of the Caribbean Reef Squid: Sepioteuthis sepioidea*. Advances in Ethology, Vol.25. Pp.150. ISBN 3-489-61936-6. (Verlag Paul Parey: 1982.) DM 48.
- PACKHAM, J.R. and HARDING, D.J.L. *Ecology of Woodland Processes*. Pp.261. Pbk ISBN 0-7131-2834-8. (Edward Arnold: 1982.) £8.95.
- PARHAM, P. and STROMINGER, J. (eds). *Histocompatibility Antigens: Structure and Function*. Receptors and Recognition Series B, Vol. 14. Pp.246. ISBN 0-412-22410-0. (Chapman & Hall: 1982.) £27.50.
- PEDGLEY, D.E. *Windborne Pests and Diseases: Meteorology of Airborne Organisms*. Pp.250. ISBN 0-85312-312-8/0-470-27516-2. (Ellis Horwood/Halsted Press: 1982.) £22.50, \$45.25.
- PENNINQ DE VRIES, F.W.T. and DJITËYE, M.A. (eds). *La Productivité des Pâturages Sahéliens*. Pp.525. ISBN 90-220-0806-1. (Centre for Agricultural Publishing and Documentation: 1982.) F.55.
- PICKLES, J.O. *An Introduction to the Physiology of Hearing*. Pp.341. Hbk ISBN 0-12-554750-1; pbk ISBN 0-12-554752-8. (Academic: 1982.) Hbk £6.95, \$13; pbk np.
- ROOT, A.I. *The ABC and XYZ of Bee Culture: An Encyclopedia Pertaining to Scientific and Practical Culture of Bees*, 38th Edn. Pp.712. ISBN 0-684-17479-0. (Scribner: 1982.) \$24.95.
- SAARINEN, E. (ed.). *Conceptual Issues in Ecology*. Pallas Paperbacks, 23. Pp.356. Pbk ISBN 90-277-1391-X. (Reidel: 1982.) Dfl. 45, \$15.95.
- SCAIFE, A. (ed.). *Plant Nutrition — 1982*, Vol. 1. and Vol. 2. Proceedings of the 9th International Plant Nutrition Colloquium held August 1982, Warwick. Pp.750. Pbk ISBN 0-85198-505-X. (Commonwealth Agricultural Bureaux: 1982.) £22.
- SCHNEIDER, R.W. (ed.). *Suppressive Soils and Plant Disease*. Pp.88. Pbk ISBN 0-89054-048-9. (American Phytopathological Society: 1982.) Np.
- SHIELDS, W.M. *Philopatry, Inbreeding, and the Evolution of Sex*. Pp.245. Hbk ISBN 0-87395-617-6; pbk ISBN 0-87395-618-4. (State University of New York Press: 1982.) Hbk £36.75; pbk £12.75.
- SINGH, A. *Prince of Cats*. Pp.192. ISBN 0-224-02034-X. (Jonathan Cape: 1982.) £8.50.
- SJODIN, R.A. *Transport in Skeletal Muscle*. Pp.157. ISBN 0-471-05265-5. (Wiley: 1982.) £27.25, \$46.50.
- SMITH, A.J.E. (ed.). *Bryophyte Ecology*. Pp.511. ISBN 0-412-22340-6. (Chapman & Hall: 1982.) £35.
- SMITH, H. and GRIERSON, D. (eds). *The Molecular Biology of Plant Development*. Botanical Monographs, Vol. 18. Pp.617. ISBN 0-632-00727-3. (Blackwell Scientific Publications: 1982.) £35.
- SPIRO, T.G. (ed.). *Iron-Sulfur Proteins*. Pp.434. ISBN 0-471-07738-0. (Wiley: 1982.) £62, \$106.
- STUART, A.J. *Pleistocene Vertebrates in the British Isles*. Pp.212. ISBN 0-582-30069-X. (Longman: 1982.) £16.50.
- TAYLOR PAGE, F.J. (ed.). *Field Guide to British Deer*, 3rd Edn. Pp.95. ISBN 0-632-00978-0. (Blackwell Scientific Publications: 1982.) £3.95.
- THE ROYAL SOCIETY. *Biological Sciences B*, Vol. 216. Pp.115. No ISBN. (Royal Society: 1982.) £9.25 (UK), £9.75 (elsewhere).
- WALLWORK, J.A. *Desert Soil Fauna*, Pp.296. ISBN 0-03-055306-7. (Praeger: 1982.) \$37.50.
- WILKINSON, P.C. *Chemotaxis and Inflammation*, 2nd Edn. Pp.249. ISBN 0-443-02085-X. (Churchill Livingstone: 1982.) £19.
- WORLD HEALTH ORGANIZATION. *Quality Assurance in Diagnostic Radiology. A Guide Prepared Following a Workshop held October 1980, FRG*. Pp.65. Pbk ISBN 92-4-154164-4. (WHO: 1982.) SwFr. 11.
- GEIS, M.L. *The Language of Television Advertising*. Pp.257. ISBN 0-12-278980-6. (Academic: 1982.) \$26.50.
- GRIFFIN, J.M. and TEECE, D.J. *OPEC Behavior and World Oil Prices*. Pp.231. Hbk ISBN 0-04-338-102-2; pbk ISBN 0-04-338-103-0. (George Allen & Unwin: 1982.) Hbk £15; pbk £5.95.
- HAYES, J.E., MICHIE, D. and PAO, Y-H. (eds). *Machine Intelligence 10*. Pp.576. ISBN 0-85312-413-0/0-470-27323-2. (Ellis Horwood/Halsted: 1982.) £37.50, \$75.35.
- HEYMAN, J. *The Masonry Arch*. Pp. 117. Hbk ISBN 0-85312-500-7; pbk ISBN 0-85312-501-5; US ISBN 0-470-27544-8. (Ellis Horwood/Halsted: 1982.) Hbk £16.50, \$33.15; pbk £8.50, \$17.
- HILLEL, D. (ed.). *Advances in Irrigation*, Vol. 1. Pp. 302. ISBN 0-12-024301-6. (Academic: 1982.) \$37.50.
- HUDSON, L. *Bodies of Knowledge: The Psychological Significance of the Nude in Art*. Pp. 163. ISBN 0-297-78117-0. (Weidenfeld & Nicolson: 1982.) £12.95.
- THE INSTITUTION OF CHEMICAL ENGINEERS. *The Assessment of Major Hazards*. Symposium Series No. 71. Pp. 428. ISBN 0-85295-146-9. (The Institution of Chemical Engineers: 1982.) Np.
- KLOPF, A.H. *The Hedonistic Neuron: A Theory of Memory, Learning, and Intelligence*. Pp. 138. Pbk ISBN 0-89116-202-X. (Hemisphere: 1982.) \$19.95.
- KONNER, M. *The Tangled Wing: Biological Constraints on the Human Spirit*. Pp. 543. ISBN 0-03-057062-X. (Holt, Rinehart & Winston: 1982.) \$19.95.
- LEONDES, C.T. (ed.). *Control and Dynamic Systems: Advances in Theory and Applications*, Vol. 18. Pp. 427. ISBN 0-12-012718-0. (Academic: 1982.) \$37.50.
- LITTLECHILD, S.C. and VAIDYA, K.G. *Energy Strategies for the UK*. Pp. 233. ISBN 0-04-339029-3. (George Allen & Unwin: 1982.) £15.
- LOWE, J., LEWIS, D. and ATKINS, M. *Total Environmental Control: The Economics of Cross-Media Pollution Transfers*. Pp.126. Pbk ISBN 0-08-026276-1. (Pergamon: 1982.) £10.95.
- LUCAS, J. *The Longest Flight*. Pp. 151. ISBN 0-224-01960-0. (Jonathan Cape, London: 1982.) £6.50.
- MAYOR, F. (ed.). *Scientific Research and Social Goals: Towards a New Development Model*. Pp. 226. ISBN 0-08-028118-4. (Pergamon: 1982.) £17.50, \$35.
- McCAFFREY, D.P. *OSHA and the Politics of Health Regulation*. Pp. 192. ISBN 0-306-41050-8. (Plenum: 1982.) Np.
- MERCHANT, C. *The Death of Nature: Women, Ecology, and the Scientific Revolution*. Pp. 348. ISBN 0-7045-3049-X. (Wildwood House, London: 1982.) £8.95.
- MINEAR, R.A. and KEITH, L.H. (eds). *Water Analysis*, Vol.1: Inorganic Species, Part 1. Pp. 287. ISBN 0-12-498301-4. (Academic: 1982.) \$39.50.
- NICOLSON, I. *Sputnik to Space Shuttle: 25 Years of the Space Age*. Pp. 207. ISBN 0-283-98891-6. (Sidgwick & Jackson: 1982.) £7.95.
- OBLER, L.K. and MENN, L. (eds). *Exceptional Language and Linguistic*. Pp.372. ISBN 0-12-523680-8. (Academic: 1982.) \$34.50.
- PETIT, J-P. *The Adventures of Archibald Higgins*. *Euclid Rules OK?* Pp.63. ISBN 0-7195-3983-8. (John Murray, London: 1982.) £4.95.
- PETIT, J-P. *The Adventures of Archibald Higgins*. *Flight of Fancy*. Pp.64. ISBN 0-7195-3982-X. (John Murray, London: 1982.) £4.95.
- PETIT, J-P. *The Adventures of Archibald Higgins*. *Informagic*. Pp.72. ISBN 0-7195-3981-1. (John Murray, London: 1982.) £4.95.
- RAPOPORT, E.H. *Areography: Geographical Strategies of Species*. Pp.269. ISBN 0-08-028914-2. (Pergamon: 1982.) £14.75, \$29.50.
- RECTOR, W.F. III (ed.). *Leadership in Space for Benefits on Earth*. Advances in the Astronautical Sciences, Vol.47. Proceedings of the 28th AAS Annual Conference held October 1981, San Diego. Pp.298. Hbk ISBN 0-87703-168-1; pbk ISBN 0-87703-169-X. (American Astronautical Society: 1982.) Np.
- ROSENFELD, A. and KAK, A.C. *Digital Picture Processing*, 2nd Edn. Vol.1. Pp.435. ISBN 0-12-597301-2. (Academic: 1982.) \$35.
- ROSENFELD, A. and KAK, A.C. *Digital Picture Processing*, 2nd Edn. Vol.2. Pp.349. ISBN 0-12-597301-2. (Academic: 1982.) \$29.
- ROSENSTEIN, J.G. *Linear Orderings*. Pp.487. ISBN 0-12-597680-1. (Academic: 1982.) Np.
- RUCKER, R. *Infinity and the Mind: The Science and Philosophy of the Infinite*. Pp.342. ISBN 3-7643-3034-1. (Birkhäuser Verlag: 1982.) SwFr. 35.
- SILVERSTEIN, A.M. *Pure Politics and Impure Science: The Swine Flu Affair*. Pp.176. ISBN 0-8018-2632-2. (Johns Hopkins University Press: 1982.) £8.75.
- SINDERMAN, C.J. *Winning the Games Scientists Play*. Pp.290. ISBN 0-306-41075-3. (Plenum: 1982.) Np.

General